Statistical Process Control in a Modern Production Environment

Windfeldt, Gitte Bjørg

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Statistical Process Control in a Modern Production Environment

Gitte Bjørg Windfeldt

Statistics
Novo Nordisk A/S
Denmark

&

Department of Informatics and Mathematical Modelling
Technical University of Denmark
Denmark

IMM–PHD–2010–225

Preface

This thesis is the result of research I have carried out during my time as PhD student at Novo Nordisk A/S and Department of Informatics and Mathematical Modelling at the Technical University of Denmark.

The origin of my research is a modern production environment in which a medical device is assembled. With its high frequent measuring close together in space and time and advanced measuring systems, this production environment introduces new challenges for the practitioners working with statistical process control (SPC). Being an industrial PhD[1], the purpose of my research has been to address some of these challenges. I have therefore put an extensive amount of work into understanding the production in detail and to understand the problems the engineers are facing in relation to the SPC system. I have chosen to focus on three different topics which I believe are important to both the production at hand and in general. For reasons of confidentiality little is disclosed regarding the actual production and data has been transformed when relevant.

The thesis is organized as a survey followed by four papers that are grouped into the three chosen topics:

Survey

> We begin with an introduction to SPC describing the background for the papers in the thesis. Then we elaborate on the papers, describing the motivation behind them and their statistical interest, and elaborating on

[1] An industrial PhD is a business focused PhD project where the student is employed at a company and enrolled at the university at the same time – working full time on the project.

relevant issues. Finally we give directions of future work.

Extending Phase I

The first topic is about extending the control chart setup phase known as Phase I. In the classical control chart setting, Phase I is primarily used to determine the limits of the control chart based on a sample of the process. With the proliferation of computers and statistical software, much more can be done in this phase to gain a better understanding of the process at hand. With the increasing complexity of modern productions environment this is not only helpful, but also necessary – not least to verify the assumptions of the control method chosen for monitoring.

Paper 1 is "Testing for Sphericity in Phase I Control Chart Applications" which is joint work with Søren Bisgaard (see Windfeldt and Bisgaard (2009)). The paper is published in Quality Reliability Engineering International, **25**, pp. 839 – 849, 2009.

The paper is aimed at practitioners in order to help them test the assumption of independence and variance homogeneity within subgroups. An assumption that is essential when using classical Shewhart charts.

Process Monitoring

With frequent measuring of a process all small variations in the process is registered. These small variations will be detected very quickly by usual control methods. For a highly capable process this results in a lot of alarms even though the quality is not threatened. Therefore a more flexible monitoring method – hiding any complexity from the end user – is required.

Paper 2 is "Using Predictive Risk in Process Control" which is joint work with Jean-Francois Plante (see Plante and Windfeldt (2009)). The paper is to be submitted for publication. The paper describes a new method for process monitoring. The method uses a statistical model of the process and a sliding window of data to estimate the probability that the next item will be outside specifications. The method is explored numerically and a case study is provided.

Paper 3 is "Monitoring a Bivariate Process using Predictive Risk", which is work in progress. The manuscript explores the new method presented in Plante and Windfeldt (2009) in a bivariate setting.

Missing Values

The last topic deals with a problem regarding missing values which I discovered when exploring data from the production. While discussing this with Søren Bisgaard it was brought to my attention that this was not an isolated case – it was seen from time to time when working with data from industry. I also discussed it with the production engineer and we decided that it would be relevant to explore this issue further.

Paper 4 is "Assessing the Impact of Missing Values on Quality Measures of an Industrial Process" which is joint work with Niels Væver Hartvig (see Windfeldt and Hartvig (2010)). The paper is submitted for publication. The paper is a case study on a problem of missing values in which we asses the impact of the missing values on the quality measures of the process. We also provide guidelines along with software to handle similar issues.

Søren Bisgaard deserves a special thanks for many inspirational discussions on industrial statistics and for supervising me into the area of research – it has meant a lot to me both professionally and personally. I also wish to thank him and his wife Sue Ellen for their generous hospitality when my husband and I stayed in Amherst. It has been a pleasure working with Jean-Francois Plante and I thank him for that, and for being understanding for my, at times, long response time. I have also enjoyed working with my supervisor and colleague Niels Væver Hartvig and I thank him for many interesting discussions and for taking the time to be a catalyst for my work when this was needed. I also wish to thank my university supervisor Helle Rootzén for accepting me as her PhD student, and for letting me pursue my own statistical interests.

I thank Per Vase and Birger Stjernholm Madsen who originally suggested this project. Furthermore I thank Per for sharing his knowledge and thoughts on process control with me. Also I would like to thank my colleagues at the production for sharing their knowledge and answering my many questions, especially Jacob Mosesson, Jørgen Toft, Lasse Langkjær and Søren Bøgvad Petersen – their input has been extremely helpful.

I wish to thank my colleagues at Statistics and my manager Sille Esbjerg for supporting me both professionally and personally through these past years.

Last but not least I wish to thank my husband Troels and daughter Emma for their continuous love and support.

Allerød,
June 29, 2010, *Gitte Bjørg Windfeldt.*

Contents

I

Survey

Chapter 1

Introduction to Statistical Process Control

Statistical process control (SPC) consists of methods for understanding, monitoring and improving process performance over time. In this section we will give an introduction to statistical process control. The focus will be on presenting the relevant background for the papers in the thesis. For a broader introduction to SPC the reader is referred to e.g. Montgomery (2005).

1.1 Basic Concepts

In this section we present some basic concepts in SPC. For further introduction to these concepts the reader is referred to Woodall (2000) and the subsequent discussion papers in the same issue.

1.1.1 Causes of Variability and Statistical Control

In any process there will be some variability which consist of contributions from various sources. Two main concepts in SPC is the notion of *chance* and

assignable causes. In the literature this is also denoted *common cause* and *special cause*, respectively. The common cause variability is variability that is predictable from a statistical perspective. Assignable cause variability is variability that is not explained by common causes, i.e. shocks, disruptions, trends, increased variability that – at least in theory – can be traced to a specific cause.

A process is *in a state of statistical control* when it is in a state where the variation can be contributed to a system of chance causes that do not appear to change over time. If a change occurs, the process is said to be out of statistical control.

Traditionally this has been interpreted as the process having an underlying (normal) distribution which do not change over time. This interpretation has been generalized over the years to include, for instance, time series models (see e.g. Bisgaard and Kulahci (2005)), and variance components models (see e.g. Woodall and Thomas (1995) and Roes and Does (1995)).

1.1.2 Control Chart

A *control chart* is a visualization of a quality characteristic of a process over time. The quality characteristic is calculated based on a sample of the process and depicted versus time or sample number. Besides the values of the quality characteristic, a control chart also consist of a horizontal line which represents the average level of the characteristic, denoted *the center line*, and two horizontal lines, denoted *upper control limit* (UCL) and *lower control limit* (LCL). The control limits are chosen such that the probability of plotting inside the control limits is high when the process is in statistical control. An example of a control chart is given in Figure 1.1.

As mentioned in Woodall (2000) a main objective of the control chart is to *"distinguish between common cause variation and special causes variation to prevent overreaction and underreaction and thereby reducing variability and maintaining stability."*

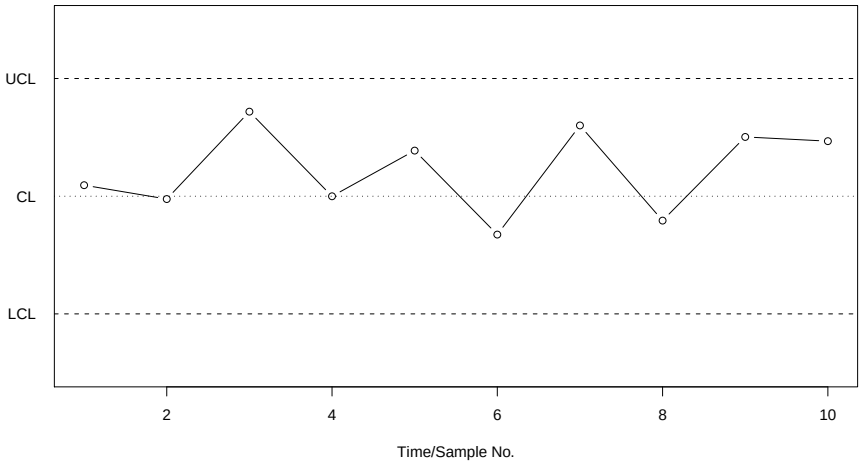

Figure 1.1: An example of a control chart.

1.1.3 Phases

Traditionally two phases are considered when working with SPC and control charts:

Phase I is a retrospective analysis on a set of historical data. The purpose of this phase is to asses whether the process was in a state of statistical control and to estimate the parameters of the underlying probability distribution. In this phase control charts are used retrospectively to determine whether the process was in statistical control.

Phase II is the monitoring phase where samples are gathered sequentially over time to maintain the process in a state of statistical control. In this phase the control chart is used prospectively to monitor the process.

The assumptions about the quality characteristics are very different in the two phases. In Phase I usually very little is assumed about the process whereas in

Phase II, the process is assumed to be in statistical control and the parameters of the process are assumed to be known or well estimated. Phase I therefore resembles exploratory and confirmatory data analysis whereas Phase II monitoring in many ways resembles hypothesis testing[1].

Control charts are used in both phases but the measure of performance of the chart is not the same. In Phase I applications measures such as the false alarm rate[2] is used to measure control chart performance. In Phase II the average run length is a common measure of control chart performance, even though the run length distribution is skewed to the right.

The majority of the SPC literature focus on control charts for Phase II monitoring of process. But as noted in Woodall (2000) it usually takes a lot of work and process understanding to get from Phase I to Phase II. This view on the importance of Phase I is shared by Thyregod and Iwersen (2000) who states: *"It is our experience that the so called Phase 1 analysis is by far the most important part of SPC. It is in this phase insight in the process and transmission of variation is obtained using the whole battery of tools in the statistical toolbox to explore the data."* In recent years there has been more focus on the pre-monitoring phase and it has been argued that the traditional view on Phase I needs to be widened. In Palm (2000) it is suggested to use three phases instead of two:

A. Charts Setup is a retrospective analysis on a set of historical data. The purpose of this stage is to get trial control limits to begin real-time monitoring. Based on historical data[3] the appropriate chart statistics are plotted together with trial control limits. If there are any signals we might consider revising the control limits with these points removed. There is not at this point in time made a judgment of whether or not the process was in statistical control. The stage could be iterative if it turns out that the data is not suitable for the purpose.

B. Process Improvement is a prospectively phase where the process is improved. As samples comes in from the process the control chart statistic is plotted on the chart with trial limits. When signal occurs they are investigated and if the cause is found it is sought removed through some

[1] There is some disagreement about these issues in the literature, see Woodall (2000).

[2] The probability of at least one signal from the chart given that the process is in statistical control.

[3] The data could have been gathered previously not necessarily for the purpose of a control chart setup.

form of process improvement. The control limits are revised as needed to reflect the improvements made. The phase ends when signals become rare and the process is considered to be stable.[4]

C. Process Monitoring is a monitoring phase where the process is monitored to detect assignable causes and maintain stability. In this phase the control chart is used prospectively to monitor the process.

This view emphasizes more clearly than the traditionally approach that there is a lot of work to be done on getting to know ones process before the process is in statistical control and the actual monitoring starts. A similar approach is suggested in a technical advice for practitioners given in Vining (2009).

The increased attention on the pre-monitoring phase has resulted in more focus in the literature on control charts specifically designed for use in this phase, see e.g. Chakraborti et al. (2009) and Jones-Farmer et al. (2009). A main part of Phase I is to estimate the parameters used for the subsequent monitoring. How to estimate these parameters and the influence of estimation on performance of the control chart in the monitoring phase are considered in e.g. Braun and Park (2008) and Jensen et al. (2006). Also different suggestions and guidelines for analysis to be performed in the pre-monitoring phase has appeared, see e.g. Mast and Trip (2009), Anderson and Whitcomb (2007), Bisgaard and Kulahci (2005) and Windfeldt and Bisgaard (2009).

1.2 Shewhart Control Charts

The Shewhart control charts for monitoring processes were developed by Shewhart in the 1920s. These charts are by far the most well-known and widely used control charts in industry. In this section we introduce the Shewhart control charts.

The Shewhart charts are based on the assumption that the quality characteristic X we wish to monitor is normally distributed with mean μ and variance σ^2. As long as the process operates according to this assumption it is said to be in statistical control, i.e. the common cause variability is described by σ^2. At time t we take a sample from the process of size n. It is assumed that the observations

[4] Palm notes that this stage could take years.

in the sample are independent normally distributed. We further assume that the sample is a *rational subgroup*, meaning that the only variation within the sample is caused by common causes. For the case $n = 1$ the reader is referred to Montgomery (2005).

The $\bar{x}$ chart is a control chart that monitors the mean of the process and its purpose is to catch unusual variability between samples. The charting statistic is the mean value of the sample which will be normally distributed with mean μ and variance σ^2/n. Assuming that μ and σ is known, the center line of the $\bar{x}$ chart is equal to μ. The control limits are of the form

$$LCL = \mu - k\frac{\sigma}{\sqrt{n}} \quad \text{and} \quad UCL = \mu + k\frac{\sigma}{\sqrt{n}},$$

where k is either chosen to be the $(1 - \alpha/2)$ quantile of the standard normal distribution yielding what is denoted *probability limits*, or as an integer yielding what is denoted *$k\sigma$-limits*. With probability limits the probability of plotting outside the limits, if the process is in statistical control, is α. When using *$k\sigma$-limits*, k is usually chosen to be 3, resulting in a probability of 0.0027 of being outside the control limits when in control.

The R chart is a control chart that monitors the range of the observations within a sample and its purpose is to catch unusual variability within subgroups. The charting statistic is the range of the sample,

$$R = \max(X_1, \ldots, X_n) - \min(X_1, \ldots, X_n).$$

The mean and standard deviation of R is usually expressed in terms of the mean and standard deviation of the relative range R/σ. The mean and standard deviation of this distribution is usually denoted d_2 and d_3[5] and values of d_2 and d_3 for sample sizes $2 \leq n \leq 25$ can be found in Appendix Table VI in Montgomery (2005). The center line for the \$R chart is therefore $d_2\sigma$ and 3σ-limits are

$$LCL = d_2\sigma - 3d_3\sigma \quad \text{and} \quad UCL = d_2\sigma + 3d_3\sigma.$$

The R chart has traditional been used because it is easy to calculate compared to the standard deviation but today this is usual not an issue.

[5]Suppressing that they depend on number of observations in the sample.

The charting statistic of the s chart is the sample standard deviation,

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

The distribution of $\sqrt{(n-1)}s/\sigma$ follows a χ distribution with $n-1$ degrees of freedom and the mean and standard deviation of s is therefore $c_4\sigma$ and $\sqrt{1-c_4^2}\sigma$, where $c_4 = \sqrt{2/(n-1)}\Gamma(n/2)/\Gamma((n-1)/2)$. Values of c_4 for $2 \leq n \leq 25$ can be found in Appendix Table VI in Montgomery (2005). The center line for the s chart is therefore $c_4\sigma$ and 3σ-limits are

$$LCL = c_4\sigma - 3\sqrt{1-c_4^2}\sigma \quad \text{and} \quad UCL = c_4\sigma + 3\sqrt{1-c_4^2}\sigma.$$

In practice μ and σ are rarely known and have to be estimated from the data in the pre-monitoring phase. Traditionally this is done based on at least $m = 25$ samples. An estimate of μ is the mean of the means for each sample $\widehat{\mu} = \bar{\bar{x}} = 1/m\sum_{i=1}^{m}\bar{x}_i$. An estimate of the range is the mean value of the ranges for each sample, i.e. $\bar{R} = 1/m\sum_{i=1}^{m}R_i$. The standard deviation σ can be estimated based on the range or the standard deviation of the samples. Traditionally the range has been used because of its computational simplicity, today this simplicity is usually less relevant. An unbiased estimator of σ based on the relative range is $\bar{R}/d_2$. As mentioned in Montgomery (2005), the range works well for small sample sizes $n \leq 6$. For larger values of $n \geq 10$, the range is loose efficiency compared to the standard deviation. An unbiased estimator of σ based on the sample standard deviation is $\tilde{\sigma} = 1/(m\,c_4)\sum_{i=1}^{m}s_i$.

The properties of the chart is usually derived under the assumption that the parameters are known. As mentioned in Section 1.1.3, the properties of a control chart in the monitoring phase is normally described by the average run length (ARL). In general, the average run length for a Shewhart chart when the process is in statistical control is $ARL_0 = \alpha^{-1}$. The $\bar{x}$ chart with 3σ-limits therefore has an average in control run length of $ARL_0 = 1/0.0027 = 370.4$. The ARL for the $\bar{x}$ chart when the process experiences a shift in the mean can be determined by considering the operating characteristic curve (OC-curve) which describes the probability of not detecting a shift in the first sample after the shift happened. This is also called the β-risk or type II error. Let μ be the level that the control chart is based on and let μ_1 be the new level. We assume that the variance

remains constant. The β-risk is then given by

$$P(LCL \leq \bar{X} \leq UCL | \mu_1 = \mu + k\sigma).$$

OC-curves for the $\bar{x}$ chart with known standard deviation is given in Figure 5.13 in Montgomery (2005). Based on these the out of control average run length can be calculated by $ARL = 1/(1-\beta)$, see Figure 5.15 in Montgomery (2005).

1.3 Process Performance

A quality characteristic usually has to meet some predescribed specifications. For a univariate quality characteristic these are usually given as an upper specification limit (USL) and a lower specification limit (LSL) which is the highest and lowest acceptable value of the characteristic, respectively. In some cases only a one side specification in prescribed. If the quality characteristic does not meet the specifications we say that the resulting product is nonconforming. There might also be prescribed a target value (T) of the quality characteristic which is the desired value of the characteristic.

The control charts ignores the specification limits and therefore does not say anything about the process' ability to meet the specifications. The fraction of nonconformities is a natural measure of process performance, but this value has traditionally been difficult to calculate. An alternative way of quantifying process performance is to use capability indices. Below we briefly describe the most widely used capability indices in industry and their relation to the fraction of nonconformities. For an introduction to capability indices see e.g. Montgomery (2005), Kotz and Lovelace (1998), Kotz and Johnson (2002), and Spiring et al. (2003).

The most well-known and widely used capability indices in industry are C_p and C_{pk}. They are both based on the assumption that the process is normally distributed with mean μ and variance σ^2 and are given by

$$C_p = \frac{USL - LSL}{6\,\sigma} \quad \text{and} \quad C_{pk} = \frac{\min(\mu - LSL, USL - \mu)}{3\,\sigma}.$$

We can see that the difference between C_p and C_{pk} is that C_{pk} not only considers the process variation, but also the location of the process. Assuming

that the process is perfectly centered at the midpoint of the specification intervals i.e. $\mu = (USL + LSL)/2$, the relationship between C_p and the fraction of nonconformities p is $p = 2\Phi(-3C_p)$, where Φ is the cumulative distribution function of the standard normal distribution. In general we have $p \geq 2\Phi(-3C_p)$ for all μ. The relationship between C_{pk} and the fraction of nonconformities is not one-to-one, but C_{pk} provides upper and lower bounds on the fraction of nonconformities given by

$$\Phi(-3C_{pk}) \leq p \leq 2\Phi(-3C_{pk}).$$

Using both C_p and C_{pk}, the relationship to the fraction of nonconformities is given by $p = \Phi(-3(2C_p - C_{pk})) + \Phi(-3C_{pk})$.

Another index that is also used in practice which is also based on the assumption of normality is

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}},$$

where T is the target value. The relation between the index C_{pm} and the fraction of nonconformities is described in Section 3.3 in Kotz and Lovelace (1998) and some of the results are mentioned below. We assume for convenience that $USL + LSL = 0$ and let $d = USL = -LSL$. If the target value is equal to the midpoint of the specification interval, the expected fraction of nonconformities is

$$p = \Phi\left(\frac{-d - \mu}{\sqrt{\lambda^2 - \mu^2}}\right) + 1 - \Phi\left(\frac{d - \mu}{\sqrt{\lambda^2 - \mu^2}}\right),$$

where $\lambda = d/(3C_{pm})$. Analytic studies of this function as a function of μ has shown that:

1. The function is symmetric about 0.
2. If $C_{pm} > 1/\sqrt{3}$ there is a local maximum at $\mu = 0$.
3. If $C_{pm} < 1/\sqrt{3}$ there is a local minimum at $\mu = 0$.
4. If $C_{pm} < 1/3$ then p increases with $|\mu|$.
5. If $1/3 < C_{pm} < 1/\sqrt{3}$ then p has a local maximum at $\mu \pm \mu_0$ for some $\mu_0 \neq 0$.

If the target value is not equal to the midpoint of the specification interval it is worth noting that the value of C_{pm} can increase even though the fraction of nonconformities increases.

1.4 Monitoring High Performance Processes

In this section we are going to review some of the methods suggested in the literature for monitoring high performance processes based on fraction of nonconformities or capability indices.

1.4.1 The Acceptance Chart

The acceptance chart which monitors the fraction of nonconformities using an $\bar{x}$ chart was introduced by Freund (1957). The acceptance chart is intended to be used when a process has a very high capability, meaning that the variability of the process is very low compared to the size of the specification interval. Assuming that the quality characteristic is normally distributed and the variance is known and constant, the mean is allowed to vary as long as the fraction of nonconformities is considered acceptable. The acceptance chart takes into account both the risk of type I error (rejecting a process that is at an acceptable level) and type II error (accepting a process at an unacceptable level) and it is closely related to acceptance sampling of variables with known variance. In acceptance sampling, type I error and type II error are also called the producers risk and consumers risk, respectively. To understand the design of an acceptance chart we first review some important concepts from acceptance sampling.

The *acceptable quality level* (AQL), denoted δ, is the highest fraction of nonconformities that we are willing to accept. The *acceptable process level* (APL) is the process level, denoted μ_δ, which corresponds to the AQL. Let LSL denote the lower specification limit and USL denote the upper specification limit. Under the assumption of a normal distribution with known variance σ^2 we have

$$\mu_{\delta,low} = LSL + Z_\delta \cdot \sigma \quad \text{and} \quad \mu_{\delta,up} = USL - Z_\delta \cdot \sigma,$$

where Z_δ is the $(1-\delta)$ quantile of the normal distribution. We note that only one specification limit is "active" at a time because of the high capability of the process.

Equivalently, the *rejectable quality level* (RQL), denoted γ, is the lowest fraction of nonconforming we would like to reject. The *rejectable process level* (RPL) is the process level μ_γ which corresponds to the RQL. Equivalent to above we have

$$\mu_{\gamma,low} = LSL + Z_\gamma \cdot \sigma \quad \text{and} \quad \mu_{\gamma,up} = USL - Z_\gamma \cdot \sigma,$$

under the assumption of a normal distribution with known variance σ^2.

When sampling from the process we have a risk, denoted α, of accepting a quality lower than δ (type I error), and a risk, denoted β, of rejecting a quality higher than γ (type II error).

There is essentially three ways to design an acceptance chart:

1. Specifying δ, the corresponding probability α, and the sample size n.
2. Specifying γ, the corresponding probability β, and the sample size n.
3. Specifying δ, γ, and the corresponding probabilities α and β.

The control limits of the $\bar{x}$ chart based on the first design are

$$LCL = LSL + \left(Z_\delta - \frac{Z_\alpha}{\sqrt{n}}\right)\sigma \quad \text{and} \quad UCL = USL - \left(Z_\delta - \frac{Z_\alpha}{\sqrt{n}}\right)\sigma.$$

Equivalently for the second design we get

$$LCL = LSL + \left(Z_\delta + \frac{Z_\beta}{\sqrt{n}}\right)\sigma \quad \text{and} \quad UCL = USL - \left(Z_\delta + \frac{Z_\beta}{\sqrt{n}}\right)\sigma.$$

When using the third design we choose the sample size so that the control limits from the first and second design are equal. This gives us a sample size of

$$n = \frac{Z_\alpha + Z_\beta}{Z_\delta + Z_\gamma}.$$

Especially when using design 3 it is important to consider whether a sample of size n can be considered to be a rational subgroup. Control limits based on design 1 are also called modified control limits (see Hill (1956) and Montgomery (2005)).

Note that when using the acceptance chart it is necessary to know the variance σ^2 or at least have a good estimate. It is usually recommended to assure that the variance remains constant by using an R chart or s chart.

Different optimizations and generalizations of the acceptance chart have been considered over the years. Wu (1998) considers an adaptive acceptance chart for the tool wear problem, where the samples size is adjusted depending on how close the mean is to the specification limits. Holmes and Mergen (2000) proposed to use an Exponential Weighted Moving Average chart (EWMA) for monitoring instead of the $\bar{x}$ chart. The EWMA chart takes individual observations allowing a stop/go decision after each observation. The acceptance chart for non-normal processes which can be approximated by a Burr distribution was introduced by Chou et al. (2005). The multivariate normal case is considered in Wesolowsky (1990), Wesolowsky (1992) and Steiner and Wesolowsky (1994). Their perspective is both acceptance sampling and acceptance charts. They consider a design based on specifying the acceptable quality level as well as the corresponding probabilities α and β. They determine the necessary sample size and control limits based on minimizing a sample cost function.

1.4.2 Monitoring Capability Indices

Instead of monitoring the fraction of nonconformities Spiring (1991) suggested to monitor a capability index in the presence of a systematic assignable cause.

The assumptions behind the long term variation of the process considered are not defined mathematically. Spiring states *"... assume only the existence of a systematic assignable cause possessing a reasonably predictable recurring pattern with known upper specification limit, lower specification limit and target value."* This is illustrated with a figure depicted in Figure 1.2. On the short term – within a sample – the observations from the process are assumed to be independent normally distributed, possibly with the mean following a linear trend.

This dynamic use of the capability index is different than traditional use. Traditionally the capability index is calculated when the process is considered to be stable and as long as it remains stable, the determined value of the index is a measure of the process performance.

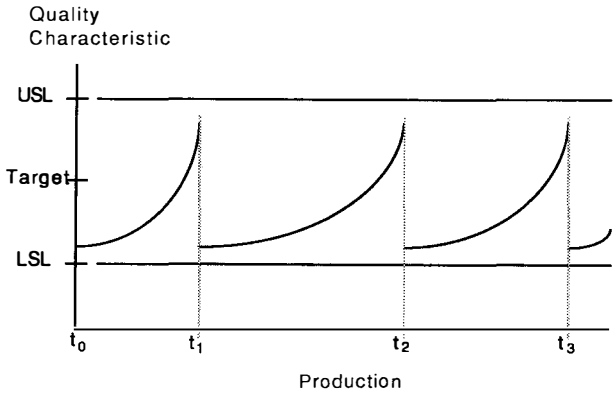

Figure 1.2: Figure 1 from Spiring (1991) illustrating a tool wear process. The values $t_0, \ldots, t_3$ are changing points (time or item number) where the process is reset.

To be able to determine the instantaneous capability of a process that is influenced by an assignable cause it is suggested to consider small time periods, e.g. a sample size n of between 5 and 25 observations where the capability can be reasonable estimated. Further it is suggested that the process capability should reflect the proximity to a target value and the variability due to random causes only[6]. This leads to determining the capability of the process at time t by a capability index of the form

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma_{rt}^2 + (\mu_t - T)^2}},$$

or

$$C_{pm}^* = \frac{\min(USL - T, T - LSL)}{3\sqrt{\sigma_{rt}^2 + (\mu_t - T)^2}},$$

where μ_t is the mean of the process at time t and σ_{rt} is the variance at time t due to random causes only. As noted in Kotz and Lovelace (1998) C_{pm}^* was introduced in Chan et al. (1988) as a generalization of C_{pm} to account for asymmetric specification limits. Note that $C_{pm}^* = C_{pm}$ if the target value is

[6]Note that the focus on the target is different from the acceptance chart where any level is good enough as long as the fraction of nonconformities is acceptable.

equal to the midpoint between the upper and lower specification limit. In Spiring (1991) both indices are denoted C_{pm}.

If the effect of the assignable cause is linear over the sample window Spiring suggests to use the variance estimate s_t^2 from a linear regression over the sample window as a measure of the random cause variation. This leads to an estimator of the capability index C_{pm}^* given by

$$\widehat{C}_{pm}^* = \frac{\min\left(USL - T, T - LSL\right)}{3\sqrt{\frac{(n-2)}{(n-1)}s_t^2 + \frac{n(\bar{x}_t - T)^2}{n-1}}}.$$

Assuming that the quality characteristic is normally distributed with mean μ and variance σ^2 the distribution of $\widehat{C}_{pm}^*$ is a non-central χ^2 distribution with $n-1$ degrees of freedom and non-centrality parameter $\lambda = n(\mu - T)^2/\sigma^2$ of the form

$$f(x) = \exp\left(-\frac{1}{2}\left(\frac{(n-1){C_{pm}^*}^2(1+\lambda/n)}{x^2} + \lambda\right)\right) \cdot \sum_{j=0}^{\infty} \frac{\left(\frac{(n-1){C_{pm}^*}^2(1+\lambda/n)}{x^2}\right)^{n-1/2+j-1}}{\Gamma\left(\frac{n-1}{2} + j\right) 2^{2j+(n-1)/2} j!},$$

see Spiring (1991). We note that the distribution of $\widehat{C}_{pm}^*$ depends on λ and C_{pm}^* (or alternatively on μ and σ).

As with the acceptance chart the process is allowed to continue as long as the capability index is above a prespecified value with a desired level of confidence. The reaction limit is based on a test of the hypothesis that $C_{pm} = C_{pm_0}^*$ and $\lambda = \lambda_0$. The value of $C_{pm_0}^*$ is directly specified as the minimum acceptable value for the capability index and λ_0 is based on specification of the sample size n and a value of $|\mu_0 - T|/\sigma_0$. The reaction limit LCL for a desired level of confidence α is then determined by the expression

$$P_{\lambda_0, C_{pm_0}}(\widehat{C}_{pm}^* < LCL) = 1 - \alpha.$$

The procedure at time t is then to calculate $\widehat{C}_{pm}^*$ based on a sample of size n and if $\widehat{C}_{pm}^* > LCL$ it is concluded that $C_{pm} > C_{pm_0}$ with confidence $1 - \alpha$.

Note that since the distribution of $\widehat{C}^*_{pm}$ depends on both λ and C^*_{pm} (or alternatively on μ and σ) the confidence level at time t will depend on the mean and variance of the process at time t. This is not elaborated on in Spiring (1991).

The hypothesis testing approach suggested in Spiring (1991) can also be used for the indices C_p and C_{pk}. The control limits can be determined from the density of the estimator and a specified minimum acceptable quality. The density of the estimators derived from the assumption of a normal distribution can be found in Kotz and Lovelace (1998). In Chou et al. (1990) the minimal value of the estimated value (the control limit in Spiring's approach) for a confidence level of 95% is calculated for various values of the minimal acceptable quality from 0.7 to 3.0 and sample sizes from 10 to 400. The hypothesis testing procedure used in Spiring (1991) was originally suggested for C_{pm} in Chan et al. (1988).

In Castagliola and Vännman (2007) is suggested a method for monitoring an unstable process by monitoring the family of capability indices $C_p(u, v)$ by an EWMA approach. The family of capability indices $C_p(u, v)$ includes the well-known indices C_p, C_{pk}, C_{pm} see Vännman (1995). It is defined by

$$C_p(u, v) = \frac{d - u|\mu - T|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}$$

where $d = (USL - LSL)/2$ and u, v are non-negative. Note that $C_p = C_p(0, 0)$, $C_{pk} = C_p(1, 0)$, and $C_{pm} = C_p(0, 1)$..

The method is a generalization of the approach for monitoring C_p suggested in Castagliola (2001). The quality characteristic is assumed to follow a normal distribution, but the process need not be stable as long as the capability of the process is constant. The assumption of a constant capability separate this method from the methods considered above.

At fixed time intervals a sample of n observations is taken and it is assumed that the observations are independent and identically normally distributed. The estimated index $\widehat{C}_p(u, v)$ is calculated based on the maximum likelihood estimates of the mean and variance of the sample for given values of u and v.

The distribution of $\widehat{C}_p(u, v)$ is skewed and depends on the values of mean value μ and variance σ of the process. To account for this, the index is transformed to obtain a variable that is approximately standard normally distributed using a two-parameter logarithm transformation of the form

$$Y_i = a_i + b_i \log(\widehat{C}_p(u, v)_i),$$

where

$$
\begin{aligned}
a_i &= -b_i \log\left(E(\widehat{C}_p(u,v)_i) \left(\frac{V(\widehat{C}_p(u,v)_i)}{E^2(\widehat{C}_p(u,v)_i)} + 1 \right)^{-1/2} \right) \\
b_i &= \left(\log\left(\frac{V(\widehat{C}_p(u,v)_i)}{E^2(\widehat{C}_p(u,v)_i)} + 1 \right) \right)^{-1/2}.
\end{aligned}
$$

Under the assumption of a constant capability index $C_p(u,v) = k_0$, a fixed set of parameters a and b are determined based on calculation of $E(\widehat{C}_p(u,v))$ and $V(\widehat{C}_p(u,v))$ using $\mu = 0$ and $\sigma = 1/(3k_0)$. The transformed variable is then monitored with an EWMA chart as described in e.g. Montgomery (2005). Assuming a known expected value and standard deviation of Y_i, the control limits are given by

$$
\begin{aligned}
LCL &= E(Y_i) - K \left(\frac{\lambda}{2-\lambda} \right)^{1/2} \sqrt{V(Y_i)}, \\
UCL &= E(Y_i) + K \left(\frac{\lambda}{2-\lambda} \right)^{1/2} \sqrt{V(Y_i)},
\end{aligned}
$$

where λ and K are constants, $0 \leq \lambda \leq 1$ and K is usually around 3.

The properties of the method is studied in Castagliola and Vännman (2008). They consider different indices and sample sizes in the range 7 to 60 under the assumption of a decrease in the capability index from 4/3 to 1 and an increase from 4/3 to 5/6. It is assumed that the process remains at the new capability level.

Chapter 2

Elaboration on the Papers

2.1 Testing for Sphericity in Phase I Control Chart Applications

The paper Windfeldt and Bisgaard (2009) is written for practitioners as a tool to test the assumption behind the $\bar{x}$ and R charts of variance homogeneity and independence of observations within subgroups. In the paper it is suggested to use the test for distributional sphericity but due to the target audience the being practitioners, the theory behind this test is not described in detail. Below we give a more detailed derivation of the test for sphericity. But first we describe the motivation and statistical interest for the paper.

2.1.1 Motivation and Statistical Interest

Classical control charts like the $\bar{x}$ and R charts are widely known and conceptual simple and therefore the natural first choice in a practical setting. Methods for setting up these charts based on a number of samples from the process have traditional been kept simple and consist mainly of calculating the control limits and plotting the points to see if they are inside these limits. But today computers and statistical software are an integrated part of the production

environment and it is therefore possible to introduce more advanced methods during Phase I. These methods can help to investigate the necessary assumptions when using control charts and help gaining a better process understanding.

The effect of non-normality has been investigated by several authors, see Montgomery (2005) for an overview, and the $\bar{x}$ chart is reasonable robust to departures from normality. Tools for checking the assumption of normality like the normal quantile plot is readily available. The assumption of temporal independence has also received a fair amount of attention in recent year and tools like the autocorrelation function are available for checking this assumption, see Montgomery (2005).

The assumption of independence and variance homogeneity within samples has received much less attention and it is this kind of violation of the assumptions we are concerned with in Windfeldt and Bisgaard (2009). A necessary condition for using the suggested method is that the observations within the sample has a consistent order that is the same across samples. It is the purpose of the $\bar{x}$ chart to catch excess variability between samples and the purpose of the R chart to catch excess variability within samples. Neither of these charts are designed to detect correlation and variance inhomogeneity within the samples. If the assumption of independence and variance homogeneity is violated the control chart may likely perform poorly in the subsequent monitoring phase, either causing excessive false alarms, not sounding valid alarms, or reacting slowly to out-of-control situations. In other words, the assumed properties of the control chart may be misleading since they are derived from the independence, equal variance and normal distribution assumptions.

The classical Shewhart charts are widely used in industry and Novo Nordisk A/S is no exception. The knowledge whether the processes violated the assumption of independence and variance homogeneity within samples is therefore of interest to Novo Nordisk A/S to be able to use the right monitoring scheme for their processes. The production at Novo Nordisk A/S range over fundamentally different types of processes. In the production involving devices, the quality characteristics are of mainly geometric dimensions. Here the knowledge that the assumption of independence and variance homogeneity within samples is violated would probably lead to an investigation and a corrective action. In the production of pharmaceuticals the quality characteristics have a different nature and the presence of correlation within the samples could very well be an inherited part of the process and in this case another monitoring scheme than

the traditional Shewhart chart might therefore be relevant.

2.1.2 The Test for Sphericity

Let $X_1, X_2, \dots X_m$ be independent multivariate normally distributed with mean μ and covariance Σ, i.e. $X_i \sim \mathcal{N}_m(\mu, \Sigma)$. We wish to test the hypothesis

$$H_0 : \Sigma = \lambda I_n \quad \text{against} \quad H_1 : \Sigma \neq \lambda I_n,$$

where $\lambda > 0$. As described in Windfeldt and Bisgaard (2009), the likelihood ratio test, also denoted the test for sphericity, rejects H_0 at a significance level of α if

$$W = \frac{\det S}{\left(\frac{1}{n}\mathrm{tr}S\right)^n} \leq k_\alpha,$$

where S is the sample covariance matrix and k_α is chosen so the significance is α.

To derive the likelihood ratio statistic we consider the likelihood function

$$L(\mu, \Sigma) = (2\pi)^{-nm/2}(\det\Sigma)^{-1/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{m}(X_i - \mu)^t\Sigma^{-1}(X_i - \mu)\right). \quad (2.1.1)$$

We have that

$$\begin{aligned}
&\sum_{i=1}^{m}(X_i-\mu)^t\Sigma^{-1}(X_i-\mu)\\
&\qquad=\sum_{i=1}^{m}(X_i-\bar{X})^t\Sigma^{-1}(X_i-\bar{X})\\
&\qquad\quad+\sum_{i=1}^{m}(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu)+2\sum_{i=1}^{m}(\bar{X}-\mu)^t\Sigma^{-1}(X_i-\bar{X})\\
&\qquad=\sum_{i=1}^{m}(X_i-\bar{X})^t\Sigma^{-1}(X_i-\bar{X})+m(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu)\\
&\qquad=\sum_{i=1}^{m}\operatorname{tr}\left((X_i-\bar{X})^t\Sigma^{-1}(X_i-\bar{X})\right)+m(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu)\\
&\qquad=\sum_{i=1}^{m}\operatorname{tr}\left(\Sigma^{-1}(X_i-\bar{X})(X_i-\bar{X})^t\right)+m(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu)\\
&\qquad=\operatorname{tr}\left(\Sigma^{-1}A\right)+m(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu),
\end{aligned}$$

where $A=(m-1)S$. We can therefore rewrite (2.1.1) as

$$\begin{aligned}
L(\mu,\Sigma)=(2\pi)^{-nm/2}&(\det\Sigma)^{-1/2}\\
&\cdot\exp\left(\operatorname{tr}\left(-\frac{1}{2}\Sigma^{-1}A\right)-\frac{m}{2}(\bar{X}-\mu)^t\Sigma^{-1}(\bar{X}-\mu)\right).
\end{aligned}$$

The likelihood ratio statistic is given by

$$\Lambda=\frac{\sup_{\mu\in\mathbb{R}^n,\,\lambda>0}\;L(\mu,\lambda I_n)}{\sup_{\mu\in\mathbb{R}^n,\,\Sigma>0}\;L(\mu,\Sigma)}.$$

The maximum for the denominator is obtained when the parameters equal the maximum likelihood estimates, i.e.

$$\widehat{\mu}=\frac{1}{m}\sum_{i=1}^{m}X_i\quad\text{and}\quad\widehat{\Sigma}=\frac{1}{m}\sum_{i=1}^{m}\left(X_i-\widehat{\mu}\right)\left(X_i-\widehat{\mu}\right)^t=\frac{1}{m}A.$$

By inserting this in the denominator we get

$$\sup_{\substack{\mu\in\mathbb{R}^n\\ \Sigma>0}}L(\mu,\Sigma)=(2\pi)^{-nm/2}m^{nm/2}\exp\left(-nm/2\right)\left(\det A\right)^{-m/2}.$$

In the numerator we get

$$\begin{aligned}\sup_{\substack{\mu\in\mathbb{R}^n\\ \lambda>0}} L(\mu,\lambda I_n) &= (2\pi)^{-nm/2} \\ &\quad \cdot \sup_{\substack{\mu\in\mathbb{R}^n\\ \lambda>0}} \left\{ \lambda^{-nm/2} \exp\left(\operatorname{tr}\left(-\frac{1}{2\lambda} A \right) + \left(-\frac{m}{2\lambda} (\bar{X}-\mu)^t (\bar{X}-\mu) \right) \right) \right\} \\ &= (2\pi)^{-nm/2} \sup_{\lambda>0} \lambda^{-nm/2} \left\{ \exp\left(\operatorname{tr}\left(-\frac{1}{2\lambda} A \right) \right) \right\} \\ &= (2\pi)^{-nm/2} \left(\frac{1}{nm} \operatorname{tr} A \right)^{-nm/2} \exp\left(-nm/2\right).\end{aligned}$$

To see the last equality note that the derivative of

$$g(\lambda) = \lambda^{-nm/2} \exp\left(\operatorname{tr}\left(-\frac{1}{2\lambda} A \right) \right)$$

is

$$g'(\lambda) = -\frac{1}{2} \exp\left(-\frac{1}{2} \frac{\operatorname{tr} A}{\lambda} \right) \lambda^{-\frac{1}{2}nm-2} \left(\lambda nm - \operatorname{tr} A \right).$$

The likelihood ratio statistic is therefore

$$\begin{aligned}\Lambda &= \frac{\left(\frac{1}{nm}\operatorname{tr} A\right)^{-nm/2} \exp\left(-nm/2\right)}{m^{nm/2} \exp\left(-nm/2\right) \left(\det A\right)^{-m/2}} \\ &= \left(\frac{\det A}{\left(\frac{1}{n}\operatorname{tr} A\right)^n} \right)^{m/2}.\end{aligned}$$

The likelihood ratio test rejects the hypothesis H_0 if the likelihood ratio statistics is small or equivalently if

$$W = \Lambda^{2/m} = \frac{\det A}{\left(\frac{1}{n}\operatorname{tr} A\right)^n} = \frac{\det S}{\left(\frac{1}{n}\operatorname{tr} S\right)^n}$$

is small.

To determine k_α we need to know the distribution of W under H_0. As noted in Muirhead (1982) the exact distribution of W is extremely complicated. An expression for the exact distribution under H_0 can be found in Nagarsenker and Pillai (1973), who also provide exact 1% and 5% quantiles for W.

According to the general likelihood theory the asymptotic distribution of $-2\log\Lambda$ is χ^2_f, where the degrees of freedom f is the number of independent parameters in the full parameter space minus the number of independent parameters under the null hypothesis, hence, $f = n(n+1)/2 - 1$. As described in Muirhead (1982) a better approximation can be found by using the work of Box (1949) and is given by

$$P\left(-2\,\frac{m-1}{m}\,\rho\log\Lambda \le X\right) = P\left(\chi^2_f \le X\right) + \omega_2\left(P\left(\chi^2_{f+4} \le X\right) - P\left(\chi^2_f \le X\right)\right),$$

where

$$\omega_2 = \frac{(n-1)(n-2)(n+2)(2n^3+6n^2+3n+2)}{288n^2(m-1)^2\rho^2},$$
$$\rho = 1 - \frac{2n^2+n+2}{6n(m-1)}.$$

The exact 5% quantiles for W and the corresponding values from the generel χ^2 approximation and the Box aproximation are compared for $n = 4, 5$, and 6 in Figures 2.1 to 2.3. We can see that the Box approximation works well for even moderate sample sizes, the differences is in the size of 10^{-4}. The general approximation is of limited use with the small sample sizes, $m \ge 25$, in the application suggested in Windfeldt and Bisgaard (2009). As can be seen from the tabulated values in Nagarsenker and Pillai (1973), the Box approximation is best for small values of n.

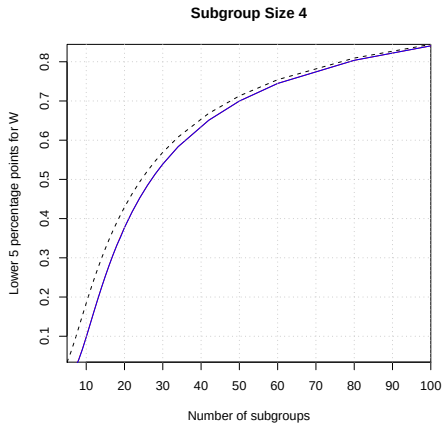

Figure 2.1: Graph of the 5% quantiles for W for $n = 4$. The solid line is the exact values while the dotted line is the generel χ^2 approximation. The box approximation is so close to the exact values that is it not visible.

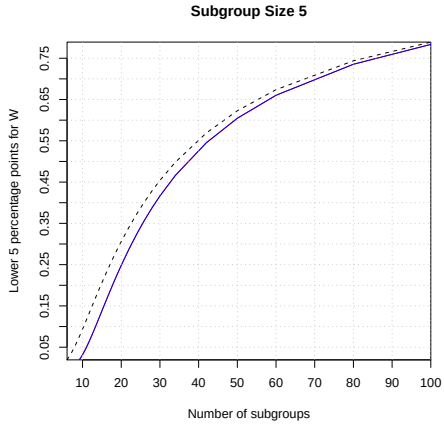

Figure 2.2: Graph of the 5% quantiles for W for $n = 5$. The solid line is the exact values while the dotted line is the generel χ^2 approximation. The box approximation is so close to the exact values that is it not visible.

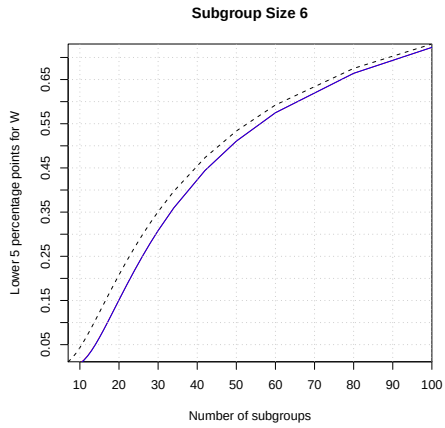

Figure 2.3: Graph of the 5% quantiles for W for $n = 6$. The solid line is the exact values while the dotted line is the generel χ^2 approximation. The box approximation is so close to the exact values that is it not visible.

2.2 Using Predictive Risk for Process Control

In this section we elaborate on different issues regarding the method for process monitoring presented in Plante and Windfeldt (2009). We begin by briefly describing the method. The reader is referred to the paper for further details. Then we describe the motivation behind the method and its statistical interest. Then we suggest ways to address the uncertainty of the estimate. A practical approach for setting up the chart is considered in Section 2.2.4. Issues regarding the sliding window approach used in the method is described in Section 2.2.5, and finally in Section 2.2.6, we relate the method to alternative methods.

2.2.1 The Method

Let $X_1, X_2, \ldots$ be a possibly multivariate characteristic that we wish to monitor for quality with specifications prescribing that the values of X must be in the set $\mathcal{S}$. As described in the paper Plante and Windfeldt (2009) we suggest to use a moving window of data to fit a parametric statistical model that estimates

the probability that the next produced item fails to meet the specifications. If the estimated probability exceeds a predetermined threshold the chart should signal. More specifically we suggest that before the t^{th} observation is collected we use a window of n data points $X_{t-n}, \ldots, X_{t-1}$ to infer the distribution $\widehat{F}_t$ for the next item to be produced. Using the inferred distribution we estimate the probability that the next item will be outside the specifications, that is $P(X_t \notin \mathcal{S})$. In the paper we suggest to use the maximum likelihood estimate $\widehat{\theta}_t$ based on the data in the window to estimate F_t. We monitor the process at time t by evaluating $1 - \int_{\mathcal{S}} dF_t(x|\widehat{\theta}_t)$ and comparing it to a predetermined threshold α.

Further we suggest to plot the estimates of the parameters from the statistical model to gain valuable insight in the process and aid in failure investigations.

2.2.2 Motivation and Statistical Interest

The motivation behind the method is found in the production environment that was the origin of my research. For a majority of the processes in this production, the sampling scheme is 100% inspection. This high frequent measuring has helped the engineers to get a very good understanding of the processes and the factors that influence them. So far the majority of the processes has been monitored on-line by monitoring the mean and variance using mainly the Shewhart $\bar{x}$ and R charts. Some modifications to the standard limits of the $\bar{x}$ chart has been made in some cases to make them operational. Various improvements have been conducted and the majority of the processes now produce very few items that do not meet the specifications. From the process engineers' perspective many of the processes are now stable in the sense that the causes of the remaining variation have been determined and deemed either impossible to remove or too resource demanding to remove from an economic and quality perspective. The processes, however, are not in statistical control. Using the control charts applied so far results in a lot of false alarms which is only natural since the processes do not live up to the stability assumptions behind the charts. Furthermore, given the high number of processes monitored, the process engineers would like to relax the surveillance on the processes with high performance and focus more on the ones that need improvement. The engineers have a high degree of process understanding and they would like to "transfer" some of this understanding into the charts, and at the same time, keep the charts simple for the operator. Since

all items produced are measured there is no risk of accepting an item that does not meet the specifications or rejecting an item that do meet specifications.

The majority of the literature on control charts for monitoring processes is about detecting special causes and maintaining stability. As described by Box and Paniagua-Quinones (2007), detection of a special cause can have three desirable results:

- The cause can be identified and eliminated,
- it may be possible to improve the process by finding the best level for the identified cause,
- the variance of the process can be reduced.

It is well recognized that using control charts can help improve a process and gain process knowledge. It is also generally accepted that when the process has been improved so it has a high performance it can be useful to relax the surveillance by using an acceptance chart and thereby allowing the process to move. As described in Section 1.4, this will allow the mean of the process to vary as long as the fraction of nonconformities is acceptable.

To use the acceptance chart the process has to be normally distributed with constant variance and a sample is assumed to consist of independent identically (normally)[1] distributed observations. Similar assumptions apply when monitoring the most well-known capability indices except the assumption of constant variance is relaxed. Traditionally the sampling from a process was small samples collected sparsely in time, making the assumptions behind the acceptance chart reasonable in many applications. Today many production environments have changed to high frequent measurements close together in space and time. Not that the traditional type of sampling does not still exist – for example in chemical laboratories where measuring can be very time consuming – but this new type of sampling is becoming more and more widespread. This does not only challenge the assumptions of the above mentioned methods but also offers an opportunity to gain and utilize valuable insight in the process and to react on different things than before. The method presented in Plante and Windfeldt (2009) is a flexible method for monitoring a wide range of processes. It is

[1] An exception to the normality assumption is considered in Chou et al. (2005).

designed for a high frequent measuring environment and allows small process instabilities with no practical importance. Further it allows the knowledge gained by the process engineer through exploratory analysis and the use of regular control charts, to be utilized in a statistical model of the process. The complexity is hidden for the operator who only has to consider one type of chart no matter the nature of the process, which can be an advantage in a production environment with many processes. The process engineer can at the same time gain valuable insight in the process by utilizing the diagnostic plots.

2.2.3 Uncertainty

The method in its current form does not take into consideration the uncertainty of the estimated parameters. One way to account for this is using bootstrap to provide a confidence interval for $\widehat{P}(X_t \notin \mathcal{S})$ rather than a point estimate. Another way would be to compensate for the uncertainty by an appropriate choice of threshold. Since the sample size is constant the sampling error would be more or less constant.

The uncertainty under a given model can be found by simulation. As an example we have used simulation to determine the uncertainty in the simple case of a normal distribution for different sample sizes. The specification limits are given by $LSL = -1$ and $USL = 1$. The mean value μ is equal to 0 and the standard deviation σ is equal to $1/6$. The result of the simulation is depicted in Figure 2.4.

2.2.4 The Setup Phase

The setup phase of the chart is not described in much detail in Plante and Windfeldt (2009). The case study section in the paper is meant to give an understanding on how such a setup could be done. The setup phase will naturally require a good cooperation between a statistician and the process engineers. A practical approach which is based on a control chart setup approach suggested by Roes and Does (1995) is given below.

1. Monitor the process intensively for a period of time.

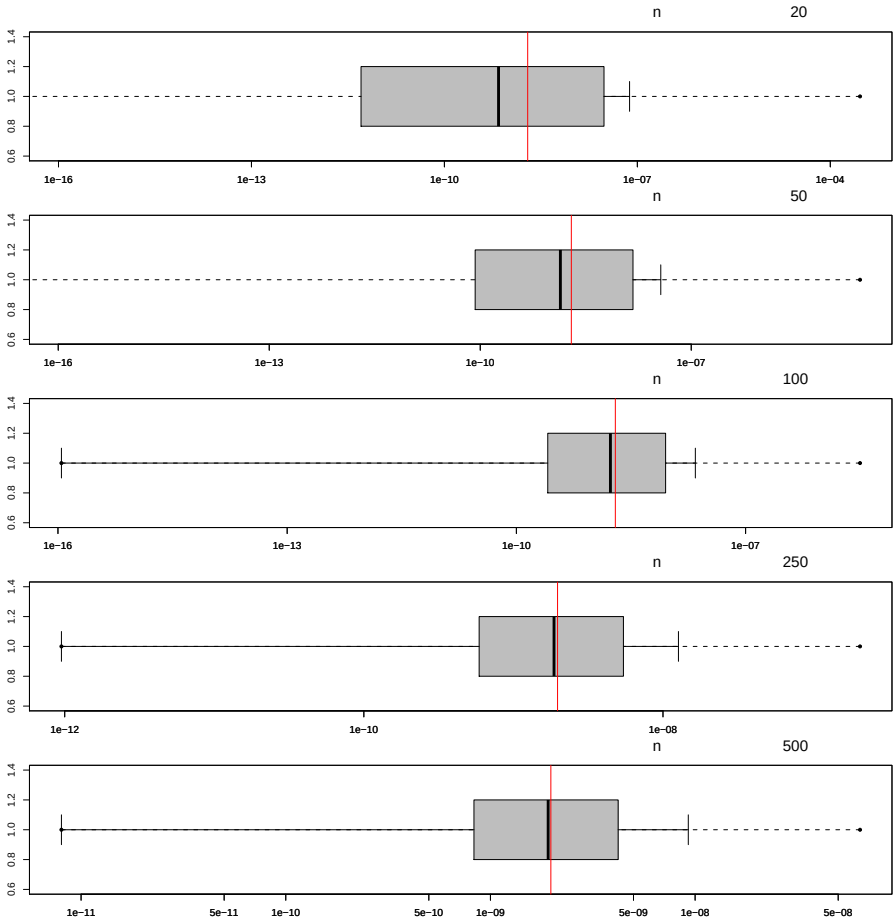

Figure 2.4: Uncertainty of $\widehat{P}(X_t \notin \mathcal{S})$ with a sample size of $n = 20, 50, 100, 250,$ and 500. The model is a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1/6$. The specification limits are $LSL = -1$ and $USL = 1$ which results in $P(X_t \notin \mathcal{S}) = 1.97^{-9}$.

2. Construct an appropriate model based on the nature of the process and supported by the data.

3. Determine which sources of variation that are removable and which are not based on practical, quality, and economical considerations. Adjust the model to account for any optimizations made to the process.

4. Determine which parameters it is necessary to estimate and which can be assumed to be fixed. Determine a good estimate for the parameters that should be fixed. Consider if there should be a diagnostic chart for the fixed parameters or another type of check to assure that they remain constant.

5. Determine an appropriate threshold for the probability of the next item not meeting specifications. This could be based on historical performance of the process or other requirements.

6. Determine an appropriate window size under consideration of the uncertainty of the estimated probability under the model. Consider if the threshold should be adjusted to account for the uncertainty.

7. Setup the chart for risk and the diagnostic plots. It is recommended to use the diagnostic charts to gain further insight in the process and assist in failure investigations.

A high performance process has often been monitored and improved by using regular control charts before it is decide to relax the surveillance and the process will therefore be well-known. This will significantly reduce the work required in the setup phase.

2.2.5 Issues Regarding the Sliding Window Approach

In Plante and Windfeldt (2009) it is suggested to use a sliding window approach. The approach means that when an abrupt change occur the data window will at some point include observations from both before and after the change. Even if one decide to use disjoint data windows the large window size would make this situation likely. When the window contains observations from both before and after the change the estimated parameters will not necessarily reflect the process after the change. There can be two possible problems with this. An

acceptable change causes a signal because the parameter estimates gives a false high probability of the next item being outside specifications. In the process used in the case study this situation occurs after periods of stand still because the machine cools down. A way to handle this situation is to flush the window and restart after longer stops.

The second problem is that an unacceptable change can fail to cause a signal because the parameters estimates gives a false low probability of the next item being outside specification. Depending on the process it can take at worst case the size of the window before the chart signals. If the process is so bad that the measured items are outside specifications it would be natural to have a mechanism that would stop the process immediately. Even if the chart does not signal the change will be visible both as an increase $\widehat{P}(X_t \in \mathcal{S})$ and a change in the diagnostic plots as the data window is filled with observations from after the change. Further the high frequent measuring that is required with this method would probably also mean that a window size of items is only a small percentage of a days production. In the process in the case study, the observations come in every three seconds and a normal days production would be around 10000 items. So given the worst case scenario, it would take approximately two and a half minute to detect the change and 50 items correspond to approximately 0.5 % of a days production.

2.2.6 Alternative Methods

Alternative methods for monitoring high performance process are described in Section 1.4. The method suggested in Plante and Windfeldt (2009) is only applicable in a high frequent measuring environment so we will restrict our attention to such an environment.

Whether to use capability indices or the fraction of nonconformities as a measure of process performance is a decision that depends on the given situation. The most well-known indices rely on the assumption of normality and independence between measurements but even when these assumptions are meet, monitoring capability indices are generally not equivalent to monitoring the fraction of nonconformities. In many industrial settings the process performance of interest is the fraction of nonconformities. Traditionally this measure has been difficult and time consuming to calculate compared to a capability index. Today the widespread use of computers and statistical software has removed this difficulty.

The values of the capability indices is though still more intuitive to many than the value of the fraction of nonconformities. Even though the scale of capability indices seems more intuitive and easier to communicate there is still an issue of ambiguity of the index values that is caused by the confounding of different characteristics of a process i.e. mean, standard deviation, closeness to target etc. Using the capability index C_{pm} as suggested by Spiring (1991) has the advantage of being able to consider a target value. But the properties of the method in this situation seems to need elaboration.

The acceptance chart introduced by Freund (1957) is designed to monitor the mean of a univariate process assuming that the process variance remains constant. It is further assumed that the observations within subgroups and between subgroups are independent. In this setting the theory behind the acceptance is simple and properties of the method, in terms of type I and type II error, is well-known. Some generalizations have appeared to some types of univariate non-normal processes and bivariate normal processes, still under the assumption of constant process variance and independence of observations. The method considered in Spiring (1991) for monitoring capability indices depends on the same assumptions as the acceptance chart, except the assumption of constant variance is relaxed. No generalization of this approach to non-standard processes has seemed to appear and it seems that generalizing this approach to processes other than the univariate normal with independent observations does not seem straightforward. Applying the method in Plante and Windfeldt (2009) to non-standard and/or multivariate processes is straightforward. So in terms of generalization, the method suggested in Plante and Windfeldt (2009) has an advantage. But there is still work to be done on exploring the properties of the method including the issue of uncertainty of the estimates.

2.3 Assessing the Impact of Missing Values on Quality Measures of an Industrial Process

The paper Windfeldt and Hartvig (2010) addresses a problem of missing values introduced by an advanced measuring system and the subsequent data handling. The primary tool in our solution to the problem is the EM algorithm. Below we formulate the EM algorithm in general terms and show an essential monotonicity property of the algorithm. We then consider the EM algorithm under the model of grouped normal data which is the model used in the paper. But we begin by describing the motivation behind the paper and its statistical interest.

2.3.1 Motivation and Statistical Interest

The motivation to the paper is a problem of missing values discovered during an investigation on how to improve the process control system of the production. To ensure the quality of an assembly process the height of the items after assembly are measured. The resolution of the measurement is three digits but from a visual display of the data it were discovered that not all values was represented. This was discussed among the production engineers and it was concluded that the measurements was transformed during the measuring process leading to an unintentional grouping of the data. An investigation into the problem was initiated with the purpose of finding the likely source of the problem and determining whether the quality measures of the process were seriously affected.

The paper Windfeldt and Hartvig (2010) is a case study of the problem of assessing the impact of the missing values on the quality measures of the process. It describes the problem, explain how we solved it and what the conclusion was. It does so in a way that others with a similar problem can see how we addressed the issue and use the software we developed. Since the problem we encountered is a general problem related to advanced measuring and the subsequent data handling it is relevant in many other industrial settings.

2.3.2 The EM Algorithm

The EM algorithm is an iterative approach for obtaining the maximum likelihood estimates in incomplete data problems. The algorithm utilizes the reduced complexity of the maximum likelihood estimation for the complete data.

Let Y be a random variable with density function $g_\theta(y)$, where $\theta \in \Theta$, which represents the observed incomplete data $y \in \mathcal{Y}$. Further let X be a random variable with density function $f_\theta(x)$ which represents the unobservable complete data $x \in \mathcal{X}$. The incomplete and complete data are connected by a one to many function $t : \mathcal{X} \to \mathcal{Y}$. Therefore

$$g_\theta(y) = \int_{t^{-1}(y)} f_\theta(x)dx.$$

The EM algorithm utilize the simplicity of the likelihood of the complete observation X to find the maximum likelihood estimate of the incomplete observation Y. Since the complete observation is unobservable, the expected complete log-likelihood function given the observed value y with respect to the current value of θ is maximized instead. More specifically, let θ_0 be an initial value of the parameter $\theta \in \Theta$. The $(k+1)^{th}$ iteration of the EM algorithm, is then

E-step. Calculate $E_{\theta_k}(\log L_x(\theta)|y)$.

M-step. Find $\theta_{k+1} \in \Theta$ such that

$$E_{\theta_k}(\log L_x(\theta_{k+1})|y) \geq E_{\theta_k}(\log L_x(\theta)|y) \quad \text{for all} \quad \theta \in \Theta.$$

In case of convergence of the likelihood values $\{L_y(\theta_k)\}_{k\in\mathbf{N}}$, the steps are continued until the difference $L_y(\theta_{k+1}) - L_y(\theta_k)$ is small.

An essential property of the EM algorithm is that the incomplete loglikehood function does not decrease after an iteration of the EM algorithm, i.e. $L_y(\theta_{k+1}) \geq L_y(\theta_k)$. To see this we look at the density of the incomplete observation

$$g_\theta(y) = \frac{f_\theta(x)}{h_\theta(x|y)},$$

where $h_\theta(x|y)$ is the conditional density of X given $Y = y$. It follows that

$$\log L_y(\theta) = \log L_x(\theta) - \log h_\theta(x|y). \tag{2.3.1}$$

Taking expectations with respect to the conditional distribution given y we get,

$$\log L_y(\theta) = E_{\theta_k}\left(\log L_x(\theta)|y\right) - E_{\theta_k}\left(\log h_\theta(x|y)|y\right).$$

Hence,

$$\begin{aligned}\log L_y(\theta_{k+1}) - \log L_y(\theta_k) = &\left(E_{\theta_k}\left(\log L_x(\theta_{k+1})|y\right) - E_{\theta_k}\left(\log L_x(\theta_k)|y\right)\right) \\ &- \left(E_{\theta_k}\left(\log h_{\theta_{k+1}}(x|y)|y\right) - E_{\theta_k}\left(\log h_{\theta_k}(x|y)|y\right)\right)\end{aligned}$$

The first difference is non-negative because of how θ_{k+1} is chosen in the M-step of the algorithm. For the second difference we have, for any $\theta \in \Theta$, that

$$\begin{aligned}E_{\theta_k}\left(\log h_\theta(x|y)|y\right) - E_{\theta_k}\left(\log h_{\theta_k}(x|y)|y\right) &= E_{\theta_k}\left(\log h_\theta(x|y) - \log h_{\theta_k}(x|y)|y\right) \\ &= E_{\theta_k}\left(\log\left(\frac{h_\theta(x|y)}{h_{\theta_k}(x|y)}\middle| y\right)\right) \\ &\leq \log\left(E_{\theta_k}\left(\frac{h_\theta(x|y)}{h_{\theta_k}(x|y)}\middle| y\right)\right) \\ &= \log\int h_\theta(x|y)dx \\ &= 0,\end{aligned}$$

where the inequality is a consequence of Jensen's equality and the concavity of the logarithm function.

2.3.3 The EM Algorithm for Grouped Observations

In this section we show how to use the EM algorithm on grouped normal observations.

Let X be a normally distributed random variable with mean μ and variance σ^2 and assume that the sampling space is divided into disjoint intervals I_i, where $i = 1, \ldots, r$. Assume that $X_1, \ldots, X_n$ are n independent observations of X but that only the number of observations n_i in I_i is known. The incomplete observation $y = (n_1, \ldots, n_r)$ will therefore follow a multinomial distribution with n draws from r categories where the probability of being in category i is

$$P(X \in I_i) = \int_{I_i} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx.$$

The incomplete likelihood function is therefore

$$L_Y(\mu, \sigma^2) \propto \prod_{i=1}^{r} P(X \in I_i)^{n_i}.$$

The loglikelihood function for the incomplete observations is therefore

$$l_Y(\mu, \sigma^2) \propto \sum_{i=1}^{r} n_i \log P(X \in I_i).$$

The complete observation is $x = (x_1, ..., x_n)$ and the likelihood function for the complete observation is

$$L_X(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

The loglikelihood function for the complete observation is therefore

$$l_X(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 .$$

The E-step of the $(k+1)^{th}$ iteration consists of finding the expected value of the complete loglikelihood function given y, i.e.

$$\begin{aligned} &E_{\mu_k,\sigma_k^2}(l_X(\mu, \sigma)|y) \\ &\qquad = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{r} n_i E_{\mu_k,\sigma_k^2}((X-\mu)^2|X \in I_i). \end{aligned}$$

We have

$$\begin{aligned} E_{\mu_k,\sigma_k^2}((X-\mu)^2|X \in I_i) &= E_{\mu_k,\sigma_k^2}\left((X-\mu)\,|X \in I_i\right)^2 + V_{\mu_k,\sigma_k^2}\left((X-\mu)\,|X \in I_i\right) \\ &= \left([X_i]_k - \mu\right)^2 + [X_i^2]_k - [X_i]_k^2, \end{aligned}$$

where

$$\begin{aligned} [X_i]_k &= E_{\mu_k,\sigma_k^2}\left(X|X \in I_i\right), \\ [X_i^2]_k &= E_{\mu_k,\sigma_k^2}\left(X^2|X \in I_i\right). \end{aligned}$$

The expected value of the complete loglikelihood function given y is therefore

$$E_{\mu_k,\sigma_k^2}\left(l_X\left(\mu_k,\sigma_k^2\right)|y\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{r} n_i\left(\left([X_i]_k-\mu\right)^2 + [X_i^2]_k - [X_i]_k^2\right). \tag{2.3.2}$$

In the M-step in the $(k+1)^{th}$ iteration we find μ_{k+1} and σ_{k+1} by maximizing (2.3.2) with respect to μ and σ. This is equivalent to maximizing the loglikelihood function for X except for the extra term $[X_i^2]_k - [X_i]_k^2$. This term is invariant when maximizing with respect to μ but should be taken into account when maximizing with respect to σ^2. We therefore get

$$\mu_{k+1} = \frac{1}{n}\sum_{i=1}^{r} n_i[X_i]_k,$$

$$\sigma_{k+1} = \frac{1}{n}\sum_{i=1}^{r} n_i\left(\left([X_i]_k-\mu_{k+1}\right)^2 + [X_i]^2 - [X_i]_k^2\right).$$

Chapter 3

Conclusion and Future Work

It has been inspiring to work closely together with process engineers and seeing their enthusiasm and dedication to process improvement. I have learned that the interest for applying statistical methods for improving process performance is readily present. But I have also learned that the modern production environment with its high frequency measurements and advanced measuring techniques is a challenge for the people working with SPC in practice. With the proliferation of computers and statistical software in the production, I believe that the same things that seem a challenge has given an opportunity to gain valuable insight into our processes like never before. This makes it an important task to provide the necessary tools for practitioners to investigate and understand the observations they obtain.

More specific direction for future work in relation to the work presented in the thesis is: In connection with the processes with grouped observations it could be relevant to consider using control charts specific for grouped observations, see Steiner et al. (1996) and Steiner (1998). The theoretical properties of the method presented in Plante and Windfeldt (2009) depends on the method of prediction and the model of the process. Further exploration of these properties are needed. We have considered using the MLE for prediction of the density

of the next observation, mainly because of the simplicity of the approach. As mentioned earlier this method does not take the uncertainty of the parameter estimates into consideration. Further work on the ideas to handle uncertainty described in Section 2.2.3 is needed. It could also be relevant to consider other ways to predict. Furthermore it could be interesting to explore the possibilities of doing a continuous check of the underlying model assumptions by considering the residuals of the fitted model.

II

Extending Phase I

Chapter 4

Testing for Sphericity in Phase I Control Chart Applications

When using $\overline{x}-R$ charts it is a crucial assumption that the observations within samples are independent and have common variance. However, this assumption is almost never checked. We propose to use the samples gathered during the phase I study and the test for distributional sphericity, to check this assumption. We supply a graph of the exact percentage points for the distribution of the test statistics. The test statistics can be computed with standard statistical software. Together with the graph of the exact percentage points, the test can easily be performed during a phase I study. We illustrate the method with examples.

4.1 Introduction

During phase I control chart studies – the retrospective analysis of (variables) control chart data in preparation for the prospective use of the control chart – a number of samples from the process under study are collected. Typically data used in phase I for $\overline{x}-R$ charts consist of $m \geq 25$ samples sampled over a representative period of time. Each sub-sample usually consist of around 5 individual observations. Based on these m sub-samples, methods have been developed for setting up $\overline{x}-R$ charts. These methods consist mainly of computing the control limits; see e.g. Montgomery (2005). Because these methods were developed a long time ago before the proliferation of modern computers and statistical software, an overriding concern was to keep them simple. However, in today's computing and software environment, what is considered simple has fundamentally changed. With current statistical software widely available in industry, it is possible to introduce more advanced methods during phase I. Such methods can help quality engineers gain information about the processes and allow for much more thorough investigations of the necessary assumptions that are made when using control charts.

Over the past few decades the assumption of temporal independence in control chart data has received a fair amount of attention; see e.g. Montgomery (2005) for an overview. Simple tools such as the autocorrelation function and other time series methods are readily available and frequently used for checking the assumption of temporal independence. However, the assumption of independence within sub-samples has received much less attention. A notable exception is Roes and Does (1995); see also the discussion in Sullivan et al. (1995). In Jensen et al. (2006) is provided a thorough review of the literature on the estimation of the parameters for control charts. It does discuss the effect of autocorrelation but does not provide information on the consequences of correlation within sub-samples. In Bischak and Trietsch (2007) the estimation of parameters for control charts is also considered using the false alarm rate. This measure is used to illustrate the danger of using to few subgroups for estimating the control limits. Neither the presence of autocorrelation or correlation within samples is considered.

In this article we demonstrate a relatively simple tool for testing the important assumptions of independence and common variance within subgroups. Independence and common variance are important to assure that the control charts have appropriate valid alarm and false alarm rates, and hence the required average

run length properties. With the increasing use of automatic sensor technology in industry, which allow measuring quality characteristics closer together in time and space, it is more than ever important to assure that these assumptions are approximately valid.

This paper is organized as follows: We begin by providing the necessary background and define the notation. Next we describe the test for sphericity and present a graph of the exact percentage points for the distribution of the test statistics. We then illustrate the method with two examples. We finish with some concluding remarks.

4.2 Background, Definitions and Notation

The standard procedure for setting up and using $\overline{x}-R$ charts is described, for example, in Chapter 5 of Montgomery (2005). We adopt the notation and refer to his book for details of this methodology. We will briefly describe the necessary assumptions when using $\overline{x}-R$ charts. The $\overline{x}-R$ charts are based on the assumption that the quality characteristic one wishes to monitor is normally distributed with mean μ and variance σ^2. It is assumed that the observations within and between subgroups are independent. A subgroup will typically consist of around $n = 5$ observations. The $\overline{x}-R$ charts are based on the assumption that this subgroup is a *rational subgroup*, meaning that within this subgroup of observations, the only cause of variation is chance. In practice μ and σ^2 are unknown and have to be estimated based on the samples available in phase I. Typically there are $m \geq 25$ subgroups available. These samples are tentatively assumed to be in *statistical control*, meaning that we proceed entertaining the tentative assumption that the data are independent and identically normally distributed. The usual way of checking this assumption is to plot the values of $\overline{x}$ and R for each subgroup on a chart with the initial trial control limits. If all points lies within the trial limits and there seem to be no systematic behavior, then we cautiously proceed provisionally assuming they are in statistical control.

It is the purpose of the $\overline{x}$ chart to catch unusual variability between the subgroups, and the purpose of the R chart is to catch excess variability within the subgroup. However, none of these control charts are designed to discover correlation or variance in-homogeneity among the observations within subgroups. It is this kind of violation of the assumptions we are concerned with in this article.

Of course, as Box famously stated in Box (1979) *"All models are wrong, but some are useful"*. Furthermore, not all assumptions are equally important. Indeed, $\overline{x}-R$ charts are relatively robust to the normality assumption. However, the assumption of independence and equal variance are typically more critical; see Appendix 3A, pp. 117–119 in Box et al. (2005). Thus, a major task during phase I is to check that the important assumptions are not seriously violated. Otherwise the control chart, when applied in the subsequent phase II, may likely perform poorly, either causing excessive false alarms, not sounding valid alarms, or reacting slowly to out-of-control situations. In other words, the assumed properties of the control chart will be misleading, possibly seriously so. Indeed, the operating characteristic (OC), or power, and the average run length (ARL) are usually derived from the independence, equal variance and normal distribution assumptions.

Mathematically we denote the observations in phase I as x_{ij}, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$. The index i counts the number of the rational subgroups, while the index j counts the observations within subgroups. The entire data set available in phase I is denoted as $\mathbf{X} = [x_{ij}]$, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$, which is a $m \times n$ matrix. Table 4.1 exemplifies such a data matrix.

Subgroup	S1	S2	S3	S4	S5	Subgroup	S1	S2	S3	S4	S5
1	140	143	137	134	135	15	144	142	143	135	144
2	138	143	143	145	146	16	140	132	144	145	141
3	139	133	147	148	139	17	137	137	142	143	141
4	143	141	137	138	140	18	137	142	142	145	143
5	142	142	145	135	136	19	142	142	143	140	135
6	136	144	143	136	137	20	136	142	140	139	137
7	142	147	137	142	138	21	142	144	140	138	143
8	143	137	145	137	138	22	139	146	143	140	139
9	141	142	147	140	140	23	140	145	142	139	137
10	142	137	145	140	132	24	134	147	143	141	142
11	137	147	142	137	135	25	138	145	141	137	141
12	137	146	142	142	140	26	140	145	143	144	138
13	142	142	139	141	142	27	145	145	137	138	140
14	137	145	144	137	140						

Table 4.1: Data from p. 14 of Grant and Leavenworth (1989). There are 27 subgroups and 5 observations in each subgroup, giving us a 27×5 data matrix.

We will further denote the ith row of the matrix $\mathbf{X}$ as a row vector $\mathbf{x}_i^T$. When thinking of $\mathbf{x}_i^T$ as a column vector, we write $\mathbf{x}_i$, which is the transpose of $\mathbf{x}_i^T$.

With these notational conventions we see that the n observations in the ith subgroup, a vector $\mathbf{x}_i$, is assumed to have a n-dimensional normal (Gaussian) distribution. That is, $\mathbf{x}_i \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the mean $\boldsymbol{\mu}$ is a vector of elements that are all equal, i.e. $\boldsymbol{\mu} = (\mu, \ldots, \mu)$, and the covariance $\boldsymbol{\Sigma}$, is a diagonal matrix with all diagonal elements equal, i.e. $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$. It is this latter assumption that is the focus of the remainder of this article. When the covariance matrix is diagonal, the observations within the subgroup are uncorrelated. Furthermore, if all the diagonal elements are the same, the individual observations within the subgroup have the same variance.

The above concepts have a nice intuitive geometric interpretation. For simplicity we consider the two dimensional case. If a two dimensional variable is normally distributed, the distribution (density) function will look like a bell shaped surface. In general, the equal level density contours of the bell shaped surface will be ellipses. If the marginal variables are dependent, the main axes of the ellipses will be slanted relative to the coordinate axes. This is illustrated in Figure 4.1. However, if the marginal variables are independent, the main

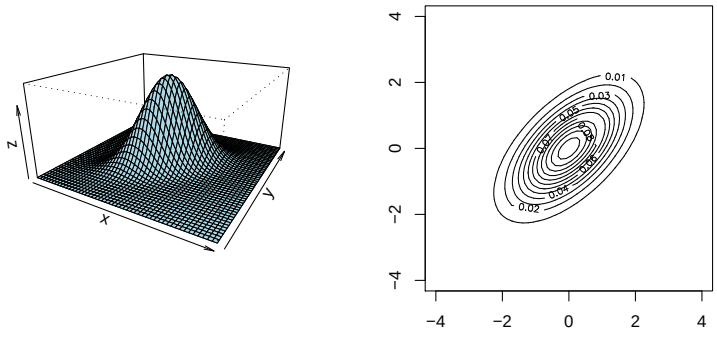

Figure 4.1: Bell shaped surface of the two dimensional normal distribution, and the corresponding contour curves, when the marginal variables are dependent. Hence, the contour curves are ellipses with the main angels slanted relative to the coordinate axes.

axes of the contour ellipses will be parallel to the coordinate axes. Furthermore

if the variances of the two marginal variables are the same, the main axes of the contour ellipses will have equal length. Hence, the ellipses will turn into circles. This is illustrated in Figure 4.2.

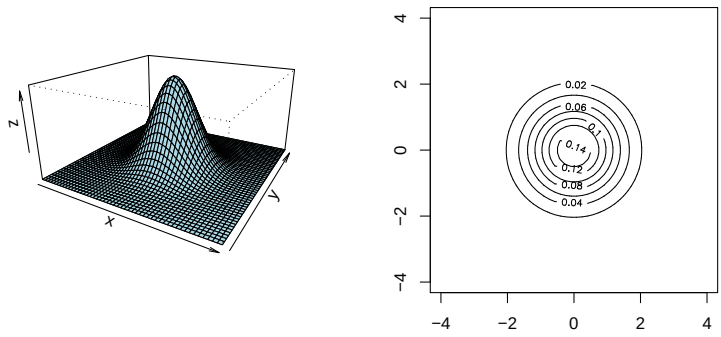

Figure 4.2: Bell shaped surface of the two dimensional normal distribution, and the corresponding contour curves, when the marginal variables are independent and have equal variance. Hence, circular contour curves.

These geometric properties carry over to any number of dimensions. The geometric implication of the above is that the k-dimensional normal distribution, as exemplified in two dimensions in Figures 4.1 and 4.2, has circular or spherical contour curves when the marginal variables are independent and have equal variance. In this case we say that the distribution of the k-dimensional variable is spherical, indeed, the distribution is rotationally symmetric. Thus, testing the assumption of independence and common variance of the observations within subgroups in the control charts setting, amounts to testing if the distribution of the subgroup vectors are spherical.

4.3 Test for Sphericity

To check if the observations within the rational subgroups are independent and have equal variance, we need to test if $\mathbf{\Sigma} = \sigma^2\mathbf{I}$. This test is based on the samples $\mathbf{x}_1, \ldots, \mathbf{x}_m$ available in phase I. Formally, the null hypothesis is $H_0 : \mathbf{\Sigma} = \sigma^2\mathbf{I}$ against the alternative $H_1 : \mathbf{\Sigma} \neq \sigma^2\mathbf{I}$. We now provide a brief description of a (likelihood ratio) test for this condition. For a detailed technical description of the test, we refer to Section 10.7 of Anderson (1958).

Assume the vector of n observations from the ith rational subgroup is multivariate normal, i.e. $\mathbf{x}_i \sim \mathcal{N}_n(\boldsymbol{\mu}, \mathbf{\Sigma})$, where $i = 1, \ldots, m$. The likelihood ratio test criterion was derived in Mauchly (1940) in the form

$$W = \frac{\det(\mathbf{S})}{(\text{tr } \mathbf{S}/n)^n},$$

where

$$\mathbf{S} = \frac{1}{m-1}\sum_{i=1}^{m}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T$$

is the sample covariance matrix. We reject H_0 if

$$W \leq k_\alpha, \tag{4.3.1}$$

where the critical value k_α is chosen such that the test size is α. Note that $\mathbf{S}$ is symmetric and therefore diagonalizable, so an intuitive way of looking at the test is

$$W = \left(\frac{(\lambda_1 \cdots \lambda_n)^{1/n}}{(\lambda_1 + \cdots + \lambda_n)/n}\right)^n, \tag{4.3.2}$$

where $\lambda_1 \leq \cdots \leq \lambda_n$ are the eigenvalues of $\mathbf{S}$. This is a power of the ratio of the geometric mean and the arithmetic mean of the eigenvalues of $\mathbf{S}$. The geometric mean cannot exceed the arithmetic mean so the likelihood ratio test, where small values of W are critical, is equivalent to testing whether the eigenvalues are equal.

This test has a nice geometrical interpretation. The shape of the contour curves of a n-dimensional multivariate normal distribution is elliptical, and the lengths of the main diagonals are inversely proportional to the eigenvalues of the covariance matrix. Thus, geometrically speaking, testing if the eigenvalues are equal is equivalent to testing if the shape of the contour curves are spherical.

Some properties worth mentioning are that the test is unbiased, see Gleser (1966) and Sugiura and Nagao (1968), and the power function is a monotone nondecreasing function of λ_k/λ_{k+1} , $1 \leq k \leq n-1$, while the remaining ratios λ_i/λ_{i+1}, $i = 1, \ldots, n-1, i \neq k$, are held fixed, see Carter and Srivastava (1977).

To find the critical value k_α used in the accept criterion we need to know the exact distribution of W. This is not a standard distribution. The exact distribution of W together with lower percentage point tables for one and five percent are given for $n = 4, \ldots, 10$ for various m in Nagarsenker and Pillai (1973). In the $\overline{x}-R$ control charts context the subgroup size is typically 5, but may occasionally be 4 or 6. Note that a subgroup size of one is also seen in practice but is not relevant for this paper. For practical $\overline{x}-R$ charts applications we therefore need only critical values for $\alpha = 0.05$ and $n = 4, 5, 6$. To make the spherical test operational we have, in Figure 4.3, produced a graph based on these general tables that will provide the necessary critical values. For values of m that are not tabulated we applied linear interpolation between the two nearest available values.

To use the test we compute the sample covariance matrix and its eigenvalues. This can be facilitated simply by using a Principal Component Analysis (PCA) routine supplied with most statistical software packages. Standard PCA output provides us with W by using equation (4.3.2). We then use the relevant curve in Figure 4.3 to determine whether to accept the hypothesis or not. According to inequality (4.3.1), we accept the hypothesis, or more appropriately phrased, fail to reject the hypothesis, if W lies above the curve, otherwise we reject. If the hypothesis is rejected a further investigation into the cause of the rejection is needed, hopefully leading to a better understanding of the process. Note that which alternative model should be subsequently entertained will depend on the given situation.

4.4 Examples

To illustrate the method we present two examples. The first is based on data from Grant (1946), but can also be found in numerous later editions of Grant and Leavenworth (1989). The data set is from a process making plastic knobs and is used as an illustration of a process presumed to be in statistical control. The second example is based on data from Roes and Does (1995) illustrating a

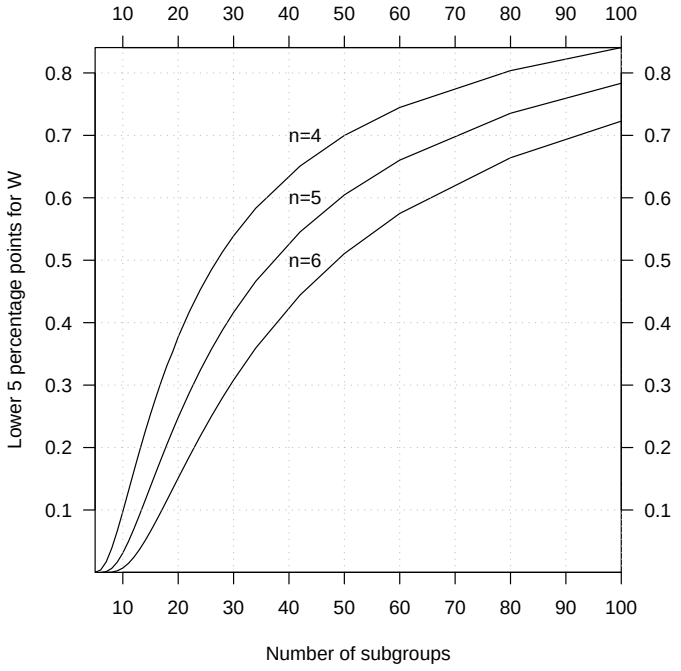

Figure 4.3: Graph of the exact lower 5 percentage points for the distribution of W for subgroup sizes $n = 4, 5$, and 6.

wafer fabrication process. This process is clearly not in statistical control.

Example 1: The data for this example have been used in numerous editions of Grant's seminal book on statistical quality control as an illustration of a process presumed to be in statistical control. The text in Grant and Leavenworth (1989) states on p. 14: "A rheostat knob, produced by plastic molding, contained a metal insert purchased from a supplier. A particular dimension determined the fit of this knob in its assembly. This dimension, which was influenced by the size of the metal insert as well as by the molding operation, was specified by the engineering department as 0.140 ± 0.003 in. Many molded knobs were rejected on 100% inspection with a go and not-go gage for failure to meet the specified tolerances. A special gage was designed and built to permit quick measurement of the actual value of this dimension. Five knobs from each hour's production were measured with this gage. Table 1-2 [Note: Same as Table 4.1] shows the measurements obtained on the first 2 days after they started."

The result of a PCA using the statistical software JMP is depicted in Table 4.2. In Figure 4.4 is a scatter plot of the observations. A scatter plot of the observations will show any obvious patterns between pairs of observations.

	S1	S2	S3	S4	S5
Eigenvalue	22.0939	13.8446	9.2173	5.6469	3.7016
Percent	40.536	25.401	16.911	10.361	6.791
Cum Percent	40.536	65.937	82.848	93.209	100.000

Table 4.2: Result of a Principal Component Analysis of the covariance matrix of the data in Table 4.1 using JMP.

We can now compute the test statistics W for the test for sphericity using the computed eigenvalues and equation (4.3.2). We get

$$W = \left(\frac{(22.0939 \cdot 13.8446 \cdot 9.2173 \cdot 5.6469 \cdot 3.7016)^{1/5}}{(22.0939 + 13.8446 + 9.2173 + 5.6469 + 3.7016)/5} \right)^5 = 0.3829.$$

In Figure 4.5 this value is inserted in the graph of the exact percentage points given in Figure 4.3. As we can see, the point lies just above the line for $n = 5$. According to inequality (4.3.1), this means that we cannot reject the hypothesis of independence and equal variance. Because the point is almost on the line for $n = 5$, it might be a good idea to investigated this issue further.

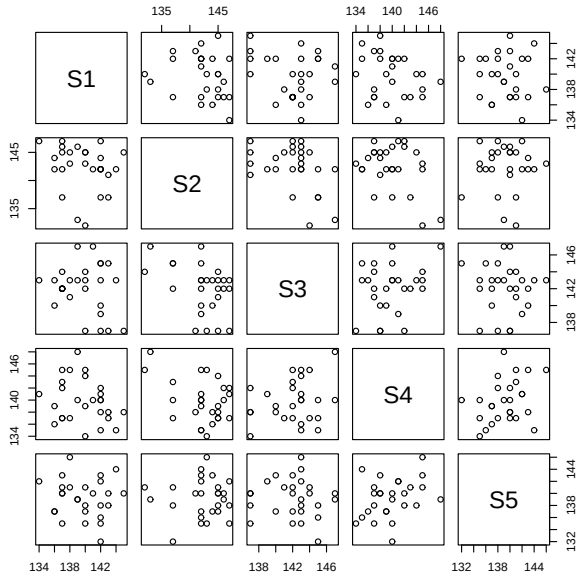

Figure 4.4: Scatter plot of the data in Table 4.1. There is no obvious pattern that indicates that the hypothesis of independence and common variance within subgroups should be rejected. The pairwise plots with $S2$ seem to have a different structure than the other plots, with a high concentration of points with high $S2$ values.

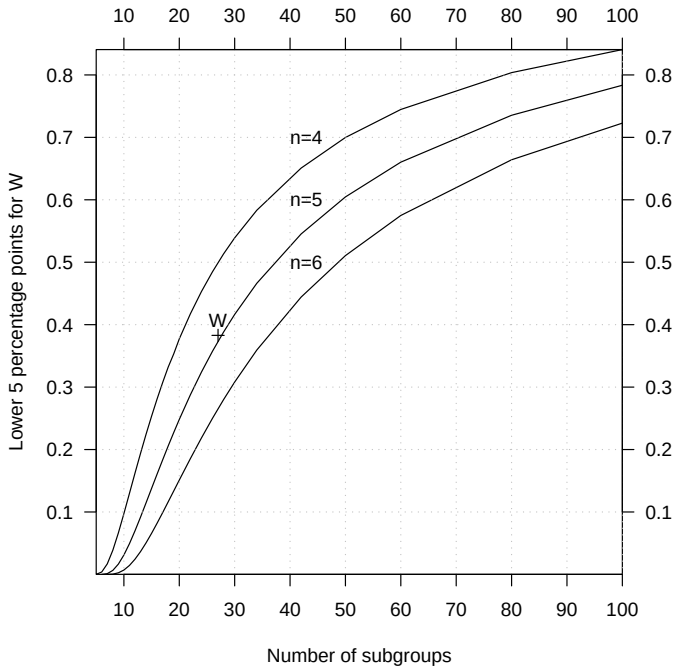

Figure 4.5: Inserting W based on the data in Table 4.1 in the graph of the exact lower five percentage points. The number of subgroups is 27.

Example 2: This data set consist of 30 batches of wafers each sampling 5 wafers from different physical locations on a surface grinder. The measurements are the thickness of the wafers after surface grinding and the unit is μm. Extreme uniformity of thickness is important. The batches are considered rational subgroups. The observations are displayed in Table 4.3. The result of a PCA using JMP is depicted in Table 4.4. A scatter plot of the observations is in Figure 4.6.

Subgroup	1	2	18	19	28	Subgroup	1	2	18	19	28
1	240	243	250	253	248	16	237	239	242	247	245
2	238	242	245	251	247	17	242	244	246	251	248
3	239	242	246	250	248	18	243	245	247	252	249
4	235	237	246	249	246	19	243	245	248	251	250
5	240	241	246	247	249	20	244	246	246	250	246
6	240	243	244	248	245	21	241	239	244	250	246
7	240	243	244	249	246	22	242	245	248	251	249
8	245	250	250	247	248	23	242	245	248	243	246
9	238	240	245	248	246	24	241	244	245	249	247
10	240	242	246	249	248	25	236	239	241	246	242
11	240	243	246	250	248	26	243	246	247	252	247
12	241	245	243	247	245	27	241	243	245	248	246
13	247	245	255	250	249	28	239	240	242	243	244
14	237	239	243	247	246	29	239	240	250	252	250
15	242	244	245	248	245	30	241	243	249	255	253

Table 4.3: Data from Roes and Does (1995). Thickness of wafers in μm. The subgroups are wafer batches and the numbers $1, 2, 18, 19$, and 28 represent physical locations on the grinder.

	1	2	18	19	28
Eigenvalue	21.2842	8.4432	3.1086	1.0874	0.8270
Percent	61.249	24.296	8.946	3.129	2.380
Cum Percent	61.249	85.545	94.491	97.620	100.00

Table 4.4: Result of a Principal Component Analysis on the covariance matrix of the data in Table 4.3 using JMP.

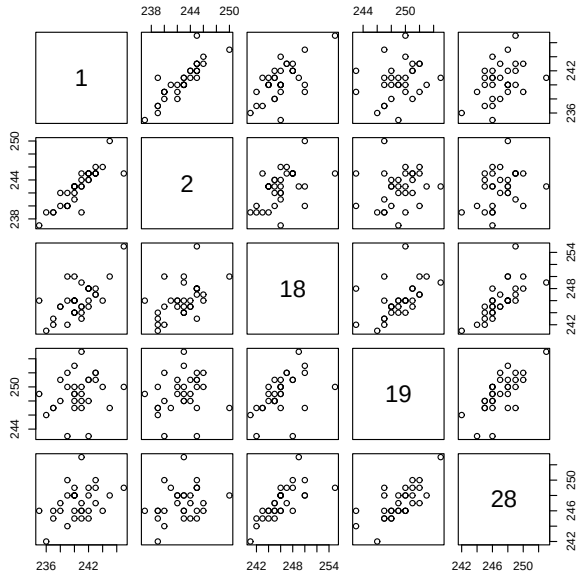

Figure 4.6: Scatter plot of the data in Table 4.3. Based on this plot we would expect, that the hypothesis of independence and common variance within subgroups will be rejected. For example, position 1 and 2 are clearly dependent.

As before, we can compute the test statistics W for the test for sphericity using the computed eigenvalues and equation (4.3.2). We get

$$W = \left(\frac{(21.2842 \cdot 8.4432 \cdot 3.1088 \cdot 1.0874 \cdot 0.8270)^{1/5}}{(21.2842 + 8.4432 + 3.1088 + 1.0874 + 0.8270)/5} \right)^5 = 0.0310.$$

In Figure 4.7 this values is inserted in the graph of the exact percentage points given in Figure 4.3. As we can see, the point lies much below the line $n = 5$,

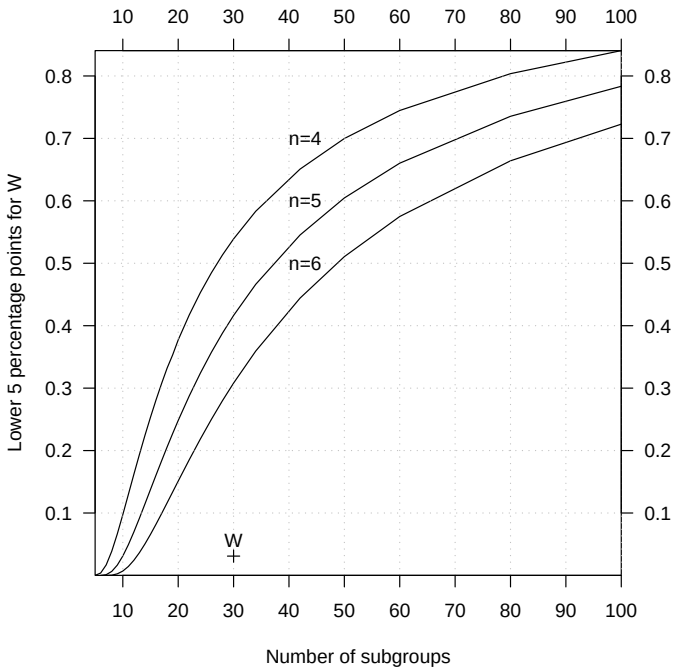

Figure 4.7: Inserting W based on the data in Table 4.3 in the graph of the exact lower five percentage points. The number of subgroups is 30.

which means we reject the hypothesis of independence and equal variance. This means that the assumptions for using $\overline{x}-R$ charts are not satisfied and therefore these should not be used directly. Instead a further investigation into the cause of this violation is needed, hopefully leading to a better understanding of the data, and subsequent leading to a useful control chart strategy. A further

investigation of the dataset used in this example and suggestions for alternative control strategies can be found in Roes and Does (1995) and the discussions following this paper.

4.5 Conclusion

As time has changed and quality engineers now have powerful computers and sophisticated statistical software readily available when they perform phase I studies (i.e. work on setting up control charts), it seems reasonable to begin to introduce more sophisticated methods for analyzing data in phase I of $\overline{x}-R$ control charts studies. This is of course not to claim that the traditional analysis using the $\overline{x}-R$ charts in phase I are obsolete or should be replaced. Our suggestion is that a number of modern graphical and statistical methods can supplement the traditional approach. To exemplify this idea we provide in this article an illustration of one such method showing how a multivariate method typically not used in the quality control context, can be be used to check if the observations in a sample are independent and have equal variance — an assumption rarely checked in the past. Indeed, these assumptions are not easily checked with the traditional control chart approach. However, if these assumptions are violated, it can potentially cause serious impairment of the performance of the control charts in phase II. Moreover, as exemplified in Roes and Does (1995) and the discussion in Sullivan et al. (1995), the knowledge that the data may be violating the standard independence and equal variance assumptions in conjunction with more focused analysis, may likely lead to important subsequent discoveries and insights relative to improving the process's performance and capability.

III

Process Monitoring

Chapter 5

Using Predictive Risk for Process Control

Classical control charts detect when a process loses its stability, but they do not give indications on the gravity of the instability and on the risk incurred if the process is not halted immediately. We suggest to use a statistical model and a window of data to evaluate the probability that the next item will not respect the specification limits. This measure of risk can then be used to decide whether the process should be halted immediately or if production could continue until it is possible to fix the process. The properties of this method are explored numerically and a case study is provided.

5.1 Introduction

Statistical process control was developed in the early 1920s. Items were typically scarcely sampled and calculations were done by hand. In this context, the simplicity of control charts was crucial and any detected instability required immediate action. Technology has evolved a lot since. Computers are ubiquitous and allow to keep track of more measurements and to use more complex models.

When large numbers of products are measured, a small variation in the product will be detected very quickly by usual control methods. Sources of instability should be fixed to improve quality, but an immediate stoppage of the production may not always be necessary. For a highly capable process, it is quite possible that changes in the process have no meaningful effect on the quality of the products. Waiting until the end of a batch, before the next shift, or until the proper parts are delivered before fixing a machine may be a better decision than stopping the production immediately.

To make good decisions on the urgency of repairs or adjustments, one needs to evaluate the risk of producing items that do not meet the quality standards described by the specifications.

We suggest to use a moving window of data to fit a statistical model that estimates the probability that the next produced item will fail to meet the specifications. When the estimated probability exceeds a predetermined threshold, production should be halted immediately for repair. Otherwise, further adjustments could be postponed, until it is more convenient to proceed.

Statistical process control involves constant efforts to reduce the variability in a process in order to keep improving the quality of the product. The proposed approach does not waive the duty of constant improvement, but it provides additional tools to make decisions on the urgency of repairs or tuning.

Modified control limits as presented by Montgomery (2005) and the acceptance control charts developed by Freund (1957) offer alternatives to the classical charts for highly capable processes. Both these methods are based on some form of specification limits, but they are best suited to monitor the mean of a process and their development beyond the normal case is limited. By comparison, the method we propose can use asymmetric or discrete models as well as linear regression or other more complex models that can include covariates in the

prediction. When multiple variables measure quality, multivariate models can be considered and will yield a single number to monitor: the probability that the next item does not meet the specifications. In all of these cases, the fitted statistical models also provide additional diagnostic tools to assist with the maintenance.

Section 2 of this paper presents the proposed methodology. Section 3 explores numerically the properties of the method in the simplified case where data are collected in rational subgroups whose size matches the window size. In Section 4, we use simulations to explore the properties of the method when the risk is evaluated after each observation is recorded. Finally, Section 5 presents a case-study based on real data.

5.2 Method

Let $X_1, X_2, \ldots$ be a possibly multivariate attribute that is monitored for quality. Specifications prescribe that the values of X must be in a given set $\mathcal{S}$, typically the interval $[LSL, USL]$ in the univariate case.

5.2.1 Concept

During the production phase, we suggest to use the control data to fit a parametric statistical model to estimate the probability that the process produces items outside of the specification limits. If this probability exceeds a pre-determined threshold, then the process should be stopped.

Using this criterion does not exclude the use of traditional charts, nor the importance to work towards a constant improvement of the process stability. However, it provides some guidance in judging if the observed instability is critical, or if it could be overlooked until a more appropriate time for maintenance.

Before the t^{th} observation is collected, we use a window of n data points, $X_{t-n}, \ldots, X_{t-1}$, to infer the distribution $\widehat{F}_t(x)$ for the next item to be produced. The probability that the item will be out of specifications, $P(X_t \notin \mathcal{S})$, is estimated using $\widehat{F}_t(x)$ and the process will be stopped for adjustments if that probability exceeds a predetermined threshold α.

In a classical setting, the ability to fulfill the required specifications is called process capability and is measured by different metrics such as C_p, C_{pk} or C_{pkm} (see Chapter 7 of Montgomery (2005) for more details). These indices are calculated when the process is considered in statistical control, i.e. yields independent and identically distributed measurements. During the production phase, classical process control focuses on stability and ignores the specification limits or the process capability. Although it is not a commonly used index, $\widehat{P}(X_t \notin \mathcal{S})$ measures the capability of a system. The proposed methodology is thus equivalent to monitoring the instantaneous capability. Such an approach is suggested by Castagliola & Vännman (2007) who develop the case where traditional measures of capability are used to monitor an univariate Gaussian characteristic collected from homogeneous subgroups.

5.2.2 Methods of Fitting

Different ways of estimating the distribution $\widehat{F}_t(x)$ can be considered. Let us first make the usual assumption that the observations are independent and identically distributed. This assumption may however be relaxed with an appropriate statistical model.

Suppose that the density function $f(x|\theta)$ and its corresponding cumulative distribution function $F(x|\theta)$ represent the process reasonably well. The maximum likelihood estimate (MLE) based on data from the window is the value of θ that maximizes $L(\theta) = \prod_{i=1}^{n} f(X_{t-i}|\theta)$, and is denoted $\widehat{\theta}_t$. We monitor the process at time t by evaluating $1 - \int_{\mathcal{S}} \mathrm{d}F(x|\widehat{\theta}_t)$ and comparing to the predetermined threshold α.

In this setting, it is also possible to weigh data relative to each others using the weighted likelihood of Hu & Zidek (2002). Allocating a larger weight to more recent data will hence yield an analogue to the Exponentially Weighted Moving Average (EWMA).

5.3 Properties of the Method in a Simplified Setting

In this section, we consider the simplified case where the data comes in rational subgroups of size n, the same size as the moving window. Therefore, every prediction is made from data that are independent of the ones previously used. Simulations in Section 5.4 study the scenario where the predictive model is fitted continuously, i.e. after the collection of each datum.

5.3.1 Normal Model

Let us suppose that $X_1, X_2, \ldots$ are normally distributed with mean μ and variance σ^2. Since the data is collected in rational subgroups, the estimates from subgroup i will be used for $t \in \{ni+1, \ldots, n(i+1)\}$. Without loss of generality, we suppose that the specification limits are $LSL = -1$ and $USL = 1$, and we denote $\mathcal{S} = [LSL, USL]$. The criterion $\widehat{P}(X_t \notin \mathcal{S}) < \alpha$ implies a set in the plane $\mu \times \sigma$ which is shown in Figure 5.1 for the cutoffs $\alpha \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-6}\}$. Values of (μ, σ) below the curves are such that $P(X_t \notin \mathcal{S}) < \alpha$.

Fitting the normal model to a sample will yield estimates of μ and σ. If $(\widehat{\mu}, \widehat{\sigma})$ falls inside the curves of Figure 5.1, the production may continue as the probability of producing faulty items is low. If $(\widehat{\mu}, \widehat{\sigma})$ is outside the curves however, it is likely that the process will produce bad items, hence it is necessary to stop immediately and proceed with tuning or maintenance.

By comparison, the usual $\bar{X}$ and S chart implies a box in the $\mu \times \sigma$ plane whose size depends on the true value of σ as well as the size of the rational subgroups. Alternatively, Chao & Cheng (1996) propose semi-circular control limits which yield better properties and depend on the same parameters. Note the essential difference: our method compares the production to the specification limits, but classical methods focus on their stability only.

Although they use specification limits, acceptance control charts are most often based on a known value of σ. They then define vertical corridors for acceptance, indifference and rejection whose sizes depend on the sample size and the selected risk levels. When the variance is estimated from the data, the vertical zones become triangles whose vertices depend on the same parameters.

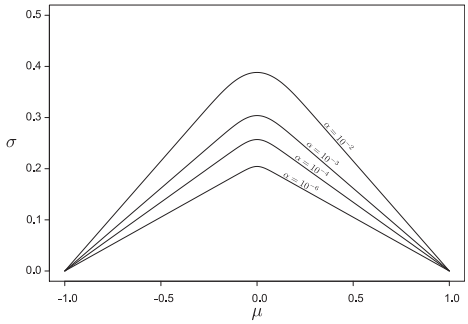

Figure 5.1: Contours implied by the criterion $P(X_t \notin [LSL, USL]) < \alpha$ for a normal response with specification limits $LSL = -1$ and $USL = 1$.

The average run length of the process under control (ARL_0) depends on the values of μ and σ under control as well as the size n of the rational subgroups.

We assume that the model is fitted once on every rational subgroup with $(\widehat{\mu}, \widehat{\sigma}) = (\bar{X}, S)$, yielding estimates that are mutually independent. Let $g(\mu)$ denote the curve from Figure 5.1 and let $\mathcal{A}$ be the zone underneath it. Production will be stopped if $(\bar{X}, S) \notin \mathcal{A}$, hence ARL_0 is the inverse of $P\{(\bar{X}, S) \notin \mathcal{A}\}$ which can be calculated numerically. Let $F_{\chi^2_{n-1}}$ be the cumulative distribution function of a χ^2_{n-1} distributed random variable and Z be a normally distributed random variable with mean μ and variance σ^2/n. Recall that $\bar{X}$ is a normal variate independent of the χ^2_{n-1} variable $(n-1)S^2/\sigma^2$. The bivariate integral to evaluate the probability that $(\bar{X}, S) \in \mathcal{A}$ can thus be approximated by

$$\begin{aligned} P\{(\bar{X}, S) \in \mathcal{A}\} &= \mathrm{E}\left[F_{\chi^2_{n-1}}\left\{\frac{n-1}{\sigma^2} g^2(Z)\right\}\right] \\ &\approx \frac{1}{N}\sum_{i=1}^{N} F_{\chi^2_{n-1}}\left\{\frac{n-1}{\sigma^2} g^2(z_i)\right\} \end{aligned} \tag{5.3.1}$$

where $z_i = \mu + (\sigma/\sqrt{n})\Phi^{-1}\{i/(N+1)\}$ are quantiles of the normal distribution and N is an arbitrary number that controls the precision of the numerical approximation. Alternatively, the expectation can also be evaluated using variants of Simpson's method.

We choose $\mu_0 = 0$ and $\sigma_0^2 = 1/36$ which yields a capability of $C_{pk} = 2$ under

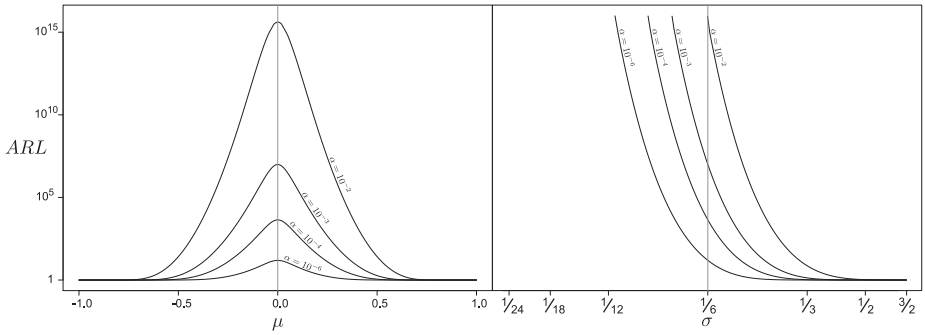

Figure 5.2: Average run length (ARL) implied by the criterion $P(X_t \notin [LSL, USL]) < \alpha$. The unit of the y-axis is the number of rational subgroups of size $n = 25$. On the left panel, different values of μ are considered with $\sigma = 1/6$. On the right panels, different values of σ are simulated while $\mu = 0$. The gray vertical lines correspond to the values under control that yield $C_{pk} = 2$, the intersections of the curve hence correspond to ARL_0. Note the logarithmic scales.

control. Figure 5.2 shows how the average run length (ARL) varies when the mean or standard error of the underlying process changes. Rational subgroups of size $n = 25$ are used; the values on the y-axis correspond to the number of such subgroups. Note how the ARL depends heavily on the risk one is willing to take. Under control, the values of ARL_0 are approximately 150, 4000, 10^6 and 10^{15} for the different α considered.

5.3.2 Weibull Model

The Weibull model is often used to model times to failure. We investigate how this model behaves when used with the proposed method. A Weibull random variable takes positive real values and its density

$$f(x) = (a/b)(x/b)^{a-1} \exp\{-(x/b)^a)\}$$

is parameterized by two real parameters, $a, b > 0$.

Let $X_1, X_2, \ldots$ be independent and identically distributed Weibull variables.

						$\log_{10}(ARL)$			
Sc.	E(X)	V(X)	a	b	$P(X \notin [1,7])$	$\alpha = 10^{-2}$	10^{-3}	10^{-4}	10^{-6}
1	4.0	0.1	15.54	4.14	2.6×10^{-10}	> 5	> 5	> 5	2.98
2	4.0	0.4	7.47	4.26	2.0×10^{-5}	> 5	2.65	0.90	0.08
3	4.4	0.1	17.16	4.54	5.4×10^{-12}	> 5	> 5	> 5	4.90
4	4.4	0.4	8.28	4.66	2.9×10^{-6}	> 5	3.98	1.72	0.24
5	6.5	0.1	25.66	6.64	2.1×10^{-2}	0.26	0.15	0.14	0.14
6	4.0	2.0	3.09	4.47	2.8×10^{-2}	0.11	0.01	0	0

Table 5.1: Characteristic of the six different scenarios that are simulated to study the ARL of the method with a Weibull distribution. The four right-most columns show the magnitude of the ARL based on 400000 simulated samples in each scenario and are expressed in numbers of subgroups of size $n = 20$.

The criterion $P(X_t \notin [LSL, USL]) < \alpha$ determines a set in the plane $a \times b$ similar to a control chart. Figure 5.3 shows these charts for different specification limits with $LSL \in \{0, 1\}$ and $USL \in \{7, \infty\}$. The different curves correspond to $\alpha \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-6}\}$.

We study the ARL of this model for rational subgroups of size $n = 20$. We suppose that $LSL = 1$ and $USL = 7$. Maximum likelihood estimates are approximately normal and that distributional property could be used to determine the ARL using an equation similar to (5.3.1). We however use another strategy and perform a Monte Carlo simulation. Six different scenarios are considered whose characteristics are summarized in Table 5.1. For each such scenario, 400000 rational subgroups of size $n = 20$ are simulated in order to estimate the probability that the MLE falls in the rejection zone. The ARL (expressed in number of rational subgroups) is then the expectation of a Geometric variable with that parameter.

All scenarios are defined based on the expectation and variance of the monitored characteristic of a single item. The appropriate equations are then solved numerically to determine the corresponding parameters. Scenario 1 is considered a well tuned process. While Scenarios 2, 3 and 4 represent a small variation in mean or variance, Scenarios 5 and 6 feature an increased mean and variance chosen such that the probability of producing an item out of the specification limits exceeds one in a thousand.

The right side of Table 5.1 shows the ARL of the proposed method under the six simulated scenarios for the different values of α. Note that > 5 occurs when

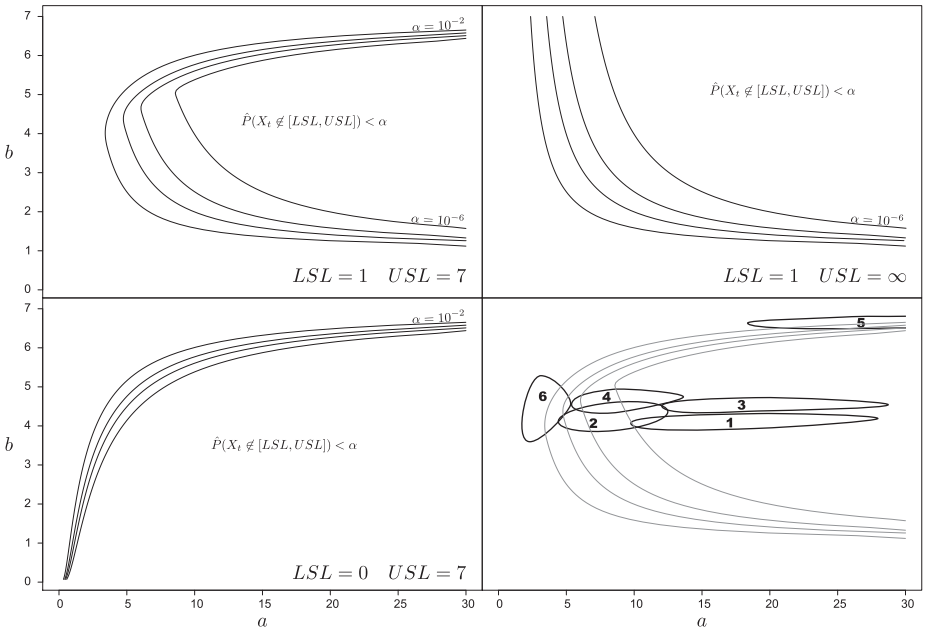

Figure 5.3: Contours of $P(X_t \notin \mathcal{S})$ for $\alpha \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-6}\}$ and different values of LSL and USL. The zone identified with $\widehat{P}(X_t \notin [LSL, USL]) < \alpha$ contains values of the parameters such that the probability of producing an item that fails to meet the specifications is bounded by α. The bottom right panel shows the area where the maximum likelihood estimates of a and b are usually found under different simulated scenarios for $n = 20$. The curves for $LSL = 1$ and $USL = 7$ appear in the same panel to put the sets in context.

none of the 400000 simulated estimates caused the method to signal a break.

The ARL found in Table 5.1 correspond to what we could expect from the lower-right panel of Figure 5.3. The MLE of the rational subgroups almost always fall in the acceptance zones for Scenarios 1 and 3. Scenarios 2 and 4 overlap with some curves for small values of α and hence should see smaller ARLs. Scenarios 5 and 6 are far enough in the rejection zone that production is stopped almost immediately.

Plots similar to Figure 5.1 and Figure 5.3 where successive MLEs are drawn could be a form of bivariate control chart. However, we believe that monitoring $\widehat{P}(X_t \notin [LSL, USL])$ is a better option.

5.4 Illustration of the Method with Continuous Evaluation

In this section, we use simulations to explore the properties of the method when the model is fitted after the collection of each datum. There is no rational subgroup in this setting and the model is based on a sliding window of n data. We explore the ARL of the method, but also simulate systems that are moving out of stability to verify if the method stops before the probability of producing an out of specification item reaches α.

5.4.1 Univariate Model

Under control, the process is assumed normal with mean 0 and unit variance. We suppose that the process meets the 6σ quality performance, hence the $\pm 6\sigma$ specification limits yielding a capability of $C_{pk} = 2$.

The maximum likelihood method is used to fit the normal model to the data. We consider two options: (i) estimating μ with $\sigma^2 = 1$ known, (ii) estimating both μ and σ^2. The threshold is set to $\alpha = 10^{-4}$. Hence, the process is stopped whenever the estimated rate of misspecified items exceeds 100 parts per million (PPM).

Production runs of at most 200000 items are simulated under different scenarios.

	μ	σ^2
Scenario A	$i/10000$	1
Scenario B	$i/1000$	1
Scenario C	0	$1 + i/10000$
Scenario D	0	$1 + i/1000$

Table 5.2: Mean and variance of the i^{th} item produced by a process losing its stability.

We record the time at which the methods signal, as well as the number of items produced out of specifications until that time.

Under control, ARL_0 is greater than 200000 when we estimate only μ since none of the runs were interrupted. When μ and σ^2 are both estimated, ARL_0 is of the order of 18000 when the moving window has size $n = 20$. When $n = 50$, ARL_0 exceeds 200000 about 99.9% of the time.

Table 5.2 summarizes the different scenarios where the products move away from their stable values. For each scenario, 100000 production runs are simulated. The run lengths are recorded and summarized by box plots in Figure 5.4. Note that the dashed lines after the whisker indicate the range of the extreme values. Due to the large number of runs, extreme values are so numerous that plotting them individually results in an almost plain line. The two curves at the bottom of the plot give the capability of the system (solid line) and the probability of producing an item out of the specification limits (dashed line). The vertical line indicates the point in time where the true probability of producing and item out of specification reaches the value of $\alpha = 10^{-4}$ that the predictive methods are aiming against. Note also the log scale for the run length.

We first note that in almost all cases, the method halts production before it gets into a state where the true probability of producing an item out of specifications reaches α. The only exception arising are in Scenario C and D where the model assuming a known variance is very slow to detect a problem due to an increasing variance. This is not a surprise.

The bulk of the stoppages are rather close to the vertical line, which means that although some early false alarms occur, they are relatively infrequent. Using a larger window size, $n = 50$ rather than $n = 20$ stabilizes the fit and makes such

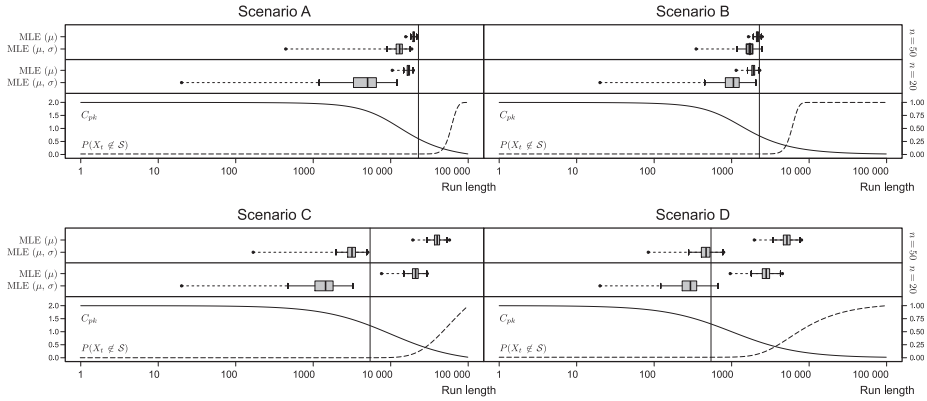

Figure 5.4: Length of 100000 runs under four scenarios where the mean or variance of items produced vary with time. The graphics at the bottom show the capability of the process and the probability that it produces an item out of specifications.

	Scenario A		Scenario B		Scenario C		Scenario D	
	$n = 20$	50	$n = 20$	50	$n = 20$	50	$n = 20$	50
MLE (μ)	1.1	3.4	2.7	6.8	10^4	10^5	10^4	10^5
MLE (μ, σ)	0.012	0.26	0.24	1.7	0.028	1.2	2.2	11

Table 5.3: Rate of items out of specification limits for the four scenarios. The figures are based on 100000 runs and are expressed in items per million.

early alarms less frequent. However, the delay in signaling may let more bad items be produced when a real problem occurs.

In the simulated runs, the number of items that did not respect the specification limits was recorded. These numbers appear in Table 5.3 in items per million. The numbers are quite reasonable, except when we use a model assuming a known variance in a scenario where the variance goes astray.

This illustrates that the proposed method allows the process to continue as long as the quality is not threatened, but sends a timely signal when such a threat appears. Before implementing standard use of this method in practical situations, similar simulations could be performed with the considered models

to ensure that they signal well under different foreseeable scenarios of instability.

5.4.2 Multivariate Model

The proposed method can be used to monitor multivariate characteristics. The purpose of the next simulation is to illustrate that, as well as the use of different likelihood methods.

Two measurements are taken on each item, yielding independent vectors $\mathbf{X}_t = (X_t, Y_t)$ assumed to follow a bivariate distribution whose margins are Gamma distributions with parameters $\alpha_x = \alpha_y = 15$ and $\beta_x = \beta_y = 3$, hence $\mathrm{E}(X_i) = \mathrm{E}(Y_i) = 5$. The two margins are linked together with the Clayton copula whose parameter θ is set such that the theoretical value of Spearman's ρ equals 0.85. Pseudo-random items are generated from this model using the algorithm of Genest & MacKay (1986) and the inverse cumulative distribution function of the Gamma distribution.

Specification limits are set to $(X, Y) \in [0.2, 15]^2$, meaning that under control, the probability of producing an item out of specification is 1.3×10^{-7} (≈ 7.7 PPM). We fix the threshold of the method to $\alpha = 10^{-3}$.

In this section we use copulas to generate the data because they are a flexible tool to generate multivariate distributions. In fact, any continuous p-dimensional distribution $F(\mathbf{x})$ can be factorized as $F(\mathbf{x}) = C\{G_1(x_1), \ldots, G_p(x_p)\}$, where G_i are the marginal distributions of F, and C, the unique copula underlying the distribution F, is a cumulative distribution function with uniform margins. The reader interested in learning more about copulas may refer to Joe (1997), Nelsen (1999), or Schweizer & Wolff (1981).

To fit the model to the data, we use the IFM method of Joe (1997) on a window of size $n = 50$. Sequentially, we

1. fit the gamma margins,

2. transform the margins into uniform distributions by plugging-in the parameters, and

3. fit the copula using the transformed margins.

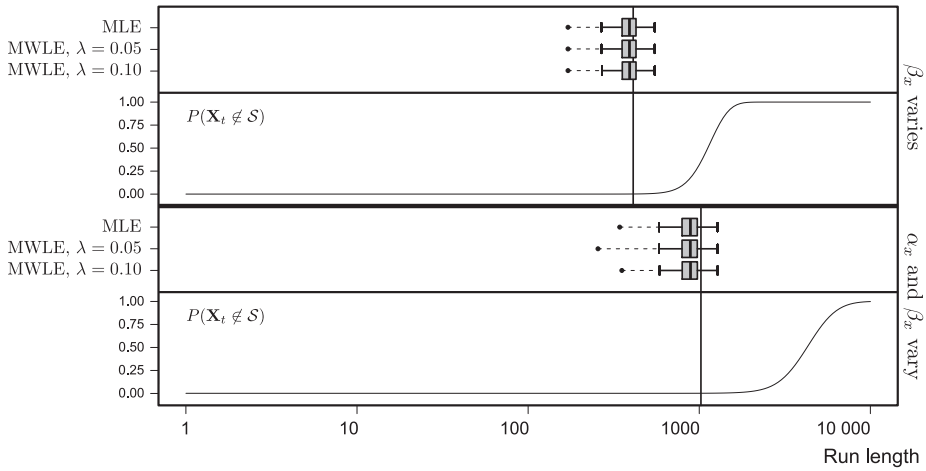

Figure 5.5: Length of 10000 production runs where the i^{th} independent measurement is a bivariate vector from a Clayton copula with Gamma margins. The graphics at the bottom show the true probability that the process produces an item out of specifications.

All fits are made using the likelihood, or the weighted likelihood of Hu & Zidek (2002) to provide a method akin to the EWMA. An intuitive justification of the weighted likelihood can also be found in Plante (2008). In short, the maximum weighted likelihood estimates in this section will maximize $WL(\theta) = \prod_{i=1}^{n} f(X_{t-i}|\theta)^{w_i}$ where $w_i \propto (1-\lambda)^i$ are positive weights that sum to 1.

The run length of a process under control is long enough to be unsuitable for a simulation, especially in a multivariate case where computation is more time consuming. We simulate different scenarios where the parameters change exponentially: for the i^{th} item, the shape parameter is $\alpha_{xi} = \alpha_x 0.999^i$ and/or the rate parameter is $\beta_{xi} = \beta_x 0.999^i$.

Figure 5.5 shows the length of 10000 runs when either β_x varies, or when α_x and β_x both vary. The vertical lines represent the time at which the process reaches a status where the probability of producing an item out of specifications reaches $\alpha = 10^{-3}$. The methods stop the processes around the right time, yielding the

Method	β_x varies	α_x and β_x vary
MLE	140	56
MWLE, $\lambda = 0.05$	51	16
MWLE, $\lambda = 0.10$	14	24

Table 5.4: Rates of items out of specifications (in PPM) for the multivariate model averaged over 10000 runs for the two different scenarios.

low rates of out of specifications items displayed in Table 5.4. There is no visible difference between the likelihood and its weighted counterparts for the run length, but the weighted method produced fewer items out of specification. A heavier weight allocated to the latest items probably allowed to blow the whistle earlier.

Although the model here is more complicated, this complexity is hidden from the user who could be monitoring a single graph showing $\widehat{P}(\mathbf{X}_t \notin \mathcal{S}) = \widehat{P}\{(X_t, Y_t) \notin [0.2, 15]^2]\}$. Figure 5.6 shows what such a graph would look like for one run under the model where both α and β vary and are estimated using the likelihood with the estimation method of Inference Functions for Margins (IFM). The upper plot shows the estimated probability of producing and item out of specifications. The lower plot is identical, except that the probit transform (the inverse of the normal cumulative distribution function) is applied to the values to make the graph more readable.

Since all data are recorded, usual control charts can be built to help the diagnostic of a problem, but additional diagnostic tools are also available. The plots of the estimated parameters in Figure 5.7, for instance, may be insightful. The parameters α_x and β_x tend to become small, while all the other parameters seem to be still bouncing around their target values indicated by the horizontal plain lines. This seems to indicate a problem with X while Y remains stable.

The estimated parameters can also be transformed into different properties of the distributions. Figure 5.8 shows the marginal means and variances for the same run. An increase in the variance of X is a likely cause for the stoppage. The two means and the variance of Y, however, look like they remained stable.

Next, we illustrate our method on an a genuine (rescaled) industrial data set.

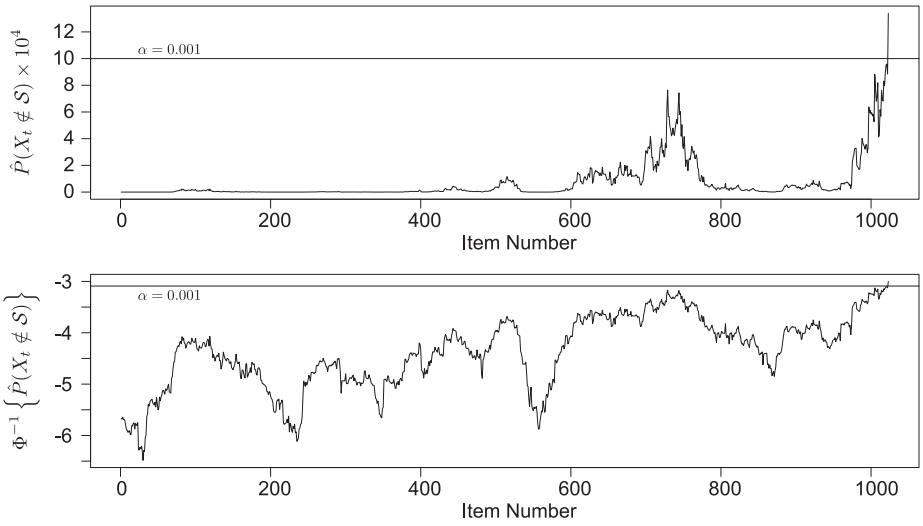

Figure 5.6: Example of graphics that an operator would monitor. The estimated probability of producing an item out of the specifications is displayed. The lower panel shows the same probability as the upper panel, but the probit transformation is applied to ease the reading.

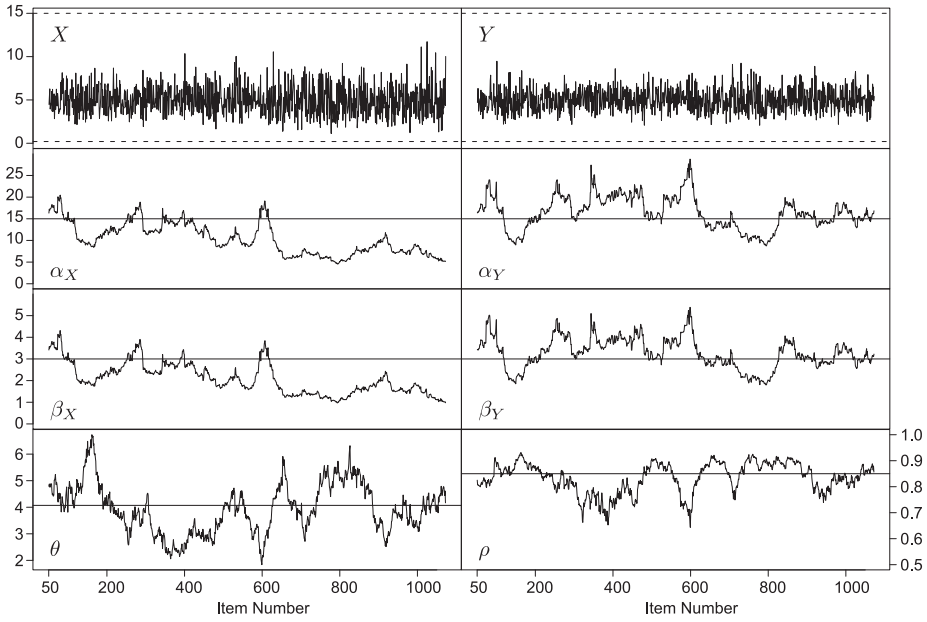

Figure 5.7: Example of a diagnostic plot showing the parameters of the statistical models estimated after each item is produced. The bivariate characteristic is simulated as a Clayton copula with Gamma margins.

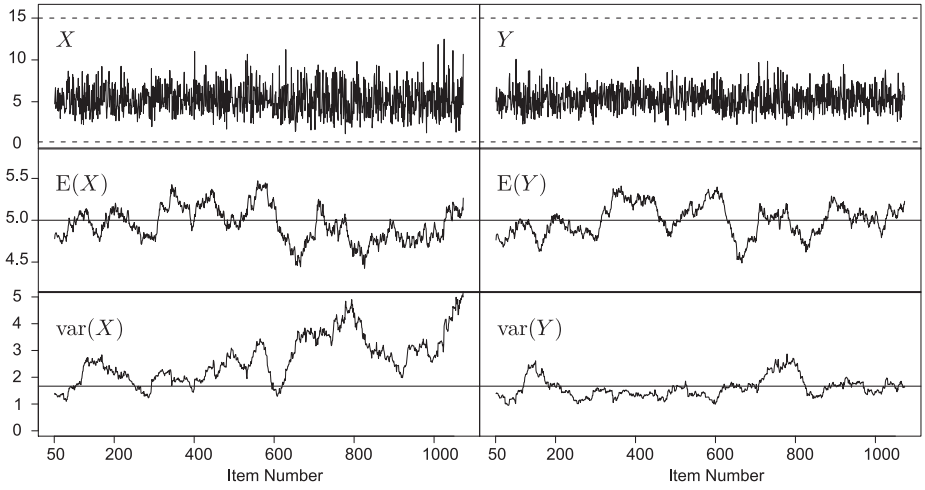

Figure 5.8: Example of a diagnostic plot showing moments of the marginal distributions based on the estimated parameters of the statistical models after each item is produced. The bivariate characteristic is simulated as a Clayton copula with Gamma margins.

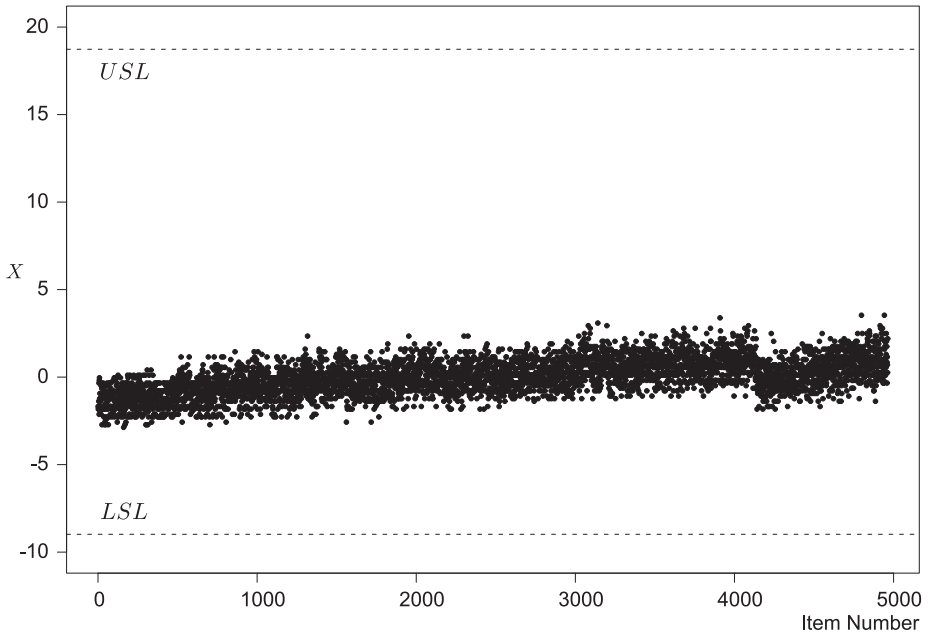

Figure 5.9: Measurements of the 4963 items produced in one day with the given machine.

5.5 A Case Study

To illustrate how to set up and use the proposed method we look at a process from the industry. The production from one machine is monitored. Each produced item is assessed by one univariate measurement X. To protect the industrial interests of Novo Nordisk A/S, the nature of the production is not discussed and the measurements are rescaled to arbitrary units.

Figure 5.9 displays the data from one day of production with the specification limits associated to the product ($LSL = -8.98$ and $USL = 18.73$). The process produces items well inside the specification limits, but it is not in statistical control.

The process experiences a drift during the day and there is a sudden drop after

Subset	Parameter	Estimate	Standard Error
1 – 4136	β_0	-1.177	0.02397
	β_1	5.297×10^{-4}	1.004×10^{-5}
	σ	0.7706	
4137 – 4963	β_0	-0.3389	0.05609
	β_1	1.801×10^{-3}	1.174×10^{-4}
	σ	0.8057	

Table 5.5: Parameter estimates for linear model fitted on two subsets of data from the day of production depicted in Figure 5.9.

4136 items. The drift is caused by the warming up of the machine during production. It is not realistic to get rid of the warm up effect on this particular machine. The abrupt change in the process is caused by a break in production, e.g. for maintenance or feeding. Such breaks occur regularly and cannot be avoided. The cooling down of the machine during the break explains the drop in the measurement of interest. Changes in components batches can affect the production likewise: the overall pattern remains the same, but the slope, level or variance may change.

Based on the engineers' experience, the data displayed in Figure 5.9 reveal a process that is stable in a practical sense: the production pattern is stable and the causes of changes have been determined and deemed too resource demanding to be dealt with. The problem is – if we were to use one of the classical control charts – that the process is not in statistical control. The drift in the process and the level changes in connection with production breaks and batch shifts would cause frequent signals, possibly resulting in tampering. By using the proposed method, we are able to handle the pattern experienced in this process and signal more wisely.

The proposed method requires choosing a model for prediction. Dividing the data in Figure 5.9 into two subsets, before and after the abrupt change, it seems natural to fit a linear model

$$X_{(\text{item } i)} = \beta_0 + \beta_1\, i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2) \tag{5.5.1}$$

on each subset. The parameter estimates are shown in Table 5.5. The usual model checking plots are depicted in Figure 5.10.

The residual plots are unstructured and the qq-plots indicate that the normality

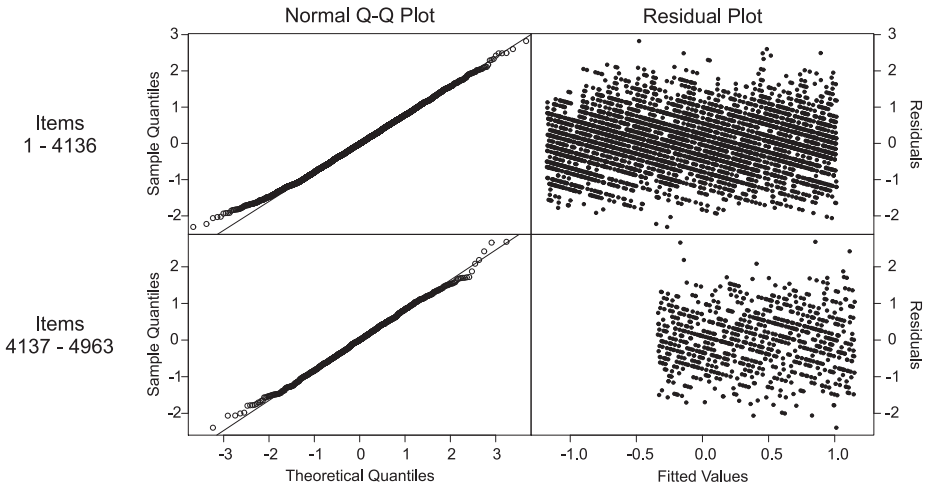

Figure 5.10: Diagnostic plots for the assumptions of the linear models on the two subsets from the day of production depicted in Figure 5.9.

assumption is reasonable. As one may guess from the small size of the standard errors relative to the parameter values, both β_0 and β_1 are significantly different from zero. Considering all of this, we conclude that the model fits well for both subsets. However, neither the intercept nor the drift estimate are equal in the two subsets. To use this model for prediction – with either item number or time as a covariate – we should estimate the model parameters continuously rather than fixing some of them to "known" values.

Item number or time may be good covariates to account for the warm-up effect, but neither can handle the cooling down happening when a stoppage occurs. If such breaks are frequent, other variables such as the temperature of the machine may be preferable and offer reliable prediction. Alternatively, the cooling down of the machine could be modeled based on the length of the stoppage – a solution that would not require investing on additional hardware.

In the absence of better covariates (e.g. the temperature of the machine), the fact that there was a stoppage (or the duration thereof) could be used to implement a strategy to account for the abrupt change. For instance, one could:

1. restart the control methods after each stoppage (i.e. wait until a window of new data is available before allowing signals), or

2. use a model that allows different intercepts before and after the stoppage. A regression model with a variable "batch number" and no intercept could be used for instance.

Some ad-hoc solutions could also be considered, e.g. fixing the estimates of β_1 and σ^2 while the first items of a batch are being produced. Remember however that the simulation results showed that fixing parameters may affect the performance of our method negatively.

The magnitude of the slopes in Table 5.5 is rather small compared to the specification limits. One could thus make the simplification that the process is normal distributed with mean μ and variance σ'^2 on a short-term basis such as a window of size 50. Empirically speaking, the normality assumption would probably remain valid. The drift of about $\sigma'/1454$ per item for the first subset and $\sigma'/447$ per item for the second subset is comparable to Scenarios 2a and 2b of Section 5.4 where the method performs well while ignoring such trend. Although parsimony would foster using the simplified model, we will fit the barely more complex regression model stated in Equation (5.5.1) for this case study.

Once the monitoring starts, the accuracy of our estimation could be affected by the abrupt change occurring after the 4136^{th} datum. The magnitude of that drop is about 1.6σ, but that value corresponds to only $0.05(USL - LSL)$. Our method might thus be able to avoid a "false alarm" (because the change does not threaten quality) where other methods would definitely halt.

After selecting a model, it is possible to decide whether the recently produced items should receive a heavier weight. This choice is akin to deciding between a moving average chart and an exponentially weighted moving average chart. For the purpose of this example, we will use the maximum likelihood with a window of size $n = 50$, hence no weighting. Implementing weighted methods would however be straightforward.

In practice, experience or regulatory demands will help for the selection of α. It would be natural to use extremely small values of α for very capable processes like the one at hand. In the presence of an incapable process, larger values of α would be relevant and will be the reflection of an increased consumer risk. Instead of choosing a specific α in this example, we plot the values of $\widehat{P}(X_t \notin \mathcal{S})$

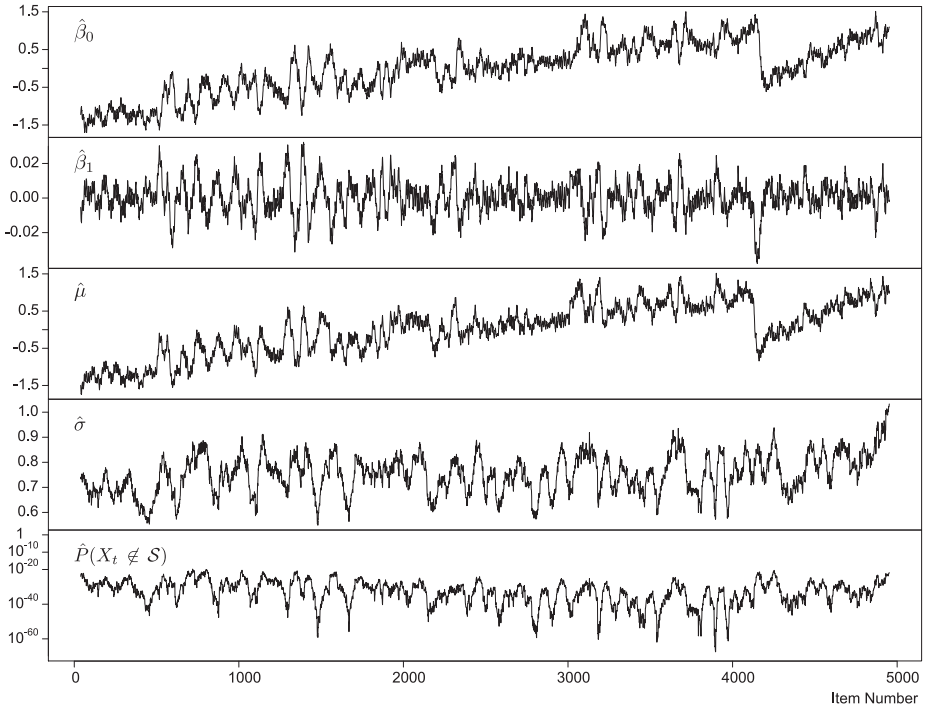

Figure 5.11: Diagnostic plots produced using our predictive method for the day of production depicted in Figure 5.9.

for Day 1. These are depicted in Figure 5.11. The behavior of the method for different values of α can then be determined from the plot. Such a plot could also be used in the setup phase, enabling the engineers to get an understanding of the capability of the process at hand.

In Figure 5.11 are also presented different panels with the diagnostic plots that can be produced from our method. Note that $\widehat{\mu}$ corresponds to the estimated mean of the next item based on β_0 and β_1. In practice, the end-user could simply monitor the last panel displaying the estimated probability of producing an item out of specification $\widehat{P}(X_t \notin \mathcal{S})$. Note that the plot for $\widehat{P}(X_t \notin \mathcal{S})$ uses a log scale to magnify the variations it undergoes. On a standard scale, the plot would look like a straight line at 0.

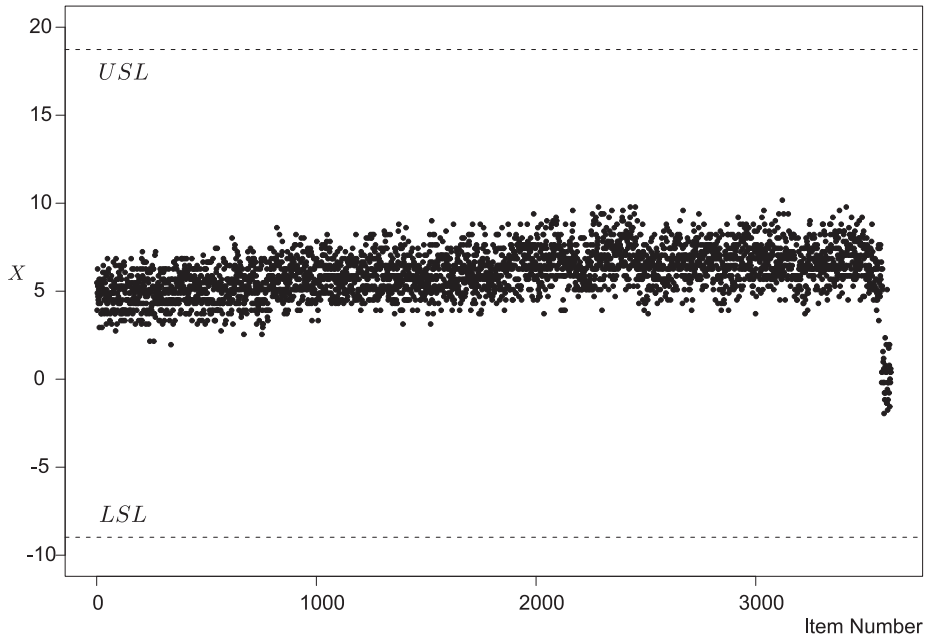

Figure 5.12: Measurements of 3615 items produced another day with the same machine. The process was halted by the engineers when he concluded that the machine had lost its stability.

For any reasonable values of α, no signal is produced during Day 1 (we would need to go as low as $\alpha = 10^{-20}$ to get one). Using a window of $n = 50$ points as the effect of smoothing the behavior of the estimates. As a consequence, the abrupt change (which is of small magnitude compared to the specification limits) does not even produce a false alarm.

Consider now a second day of data where production was stopped after an important unexplained shift in the production. Figure 5.12 displays the data from that day. As we can see, the process level falls dramatically and is then stopped by the engineer in charge. Figure 5.13 shows the diagnostic plots that are produced by our method based on the data from that day.

Everything seemed to go wrong just before the production is put to halt, the mean and variance both going astray. This trouble is picked up by $\widehat{P}(X_t \notin \mathcal{S})$

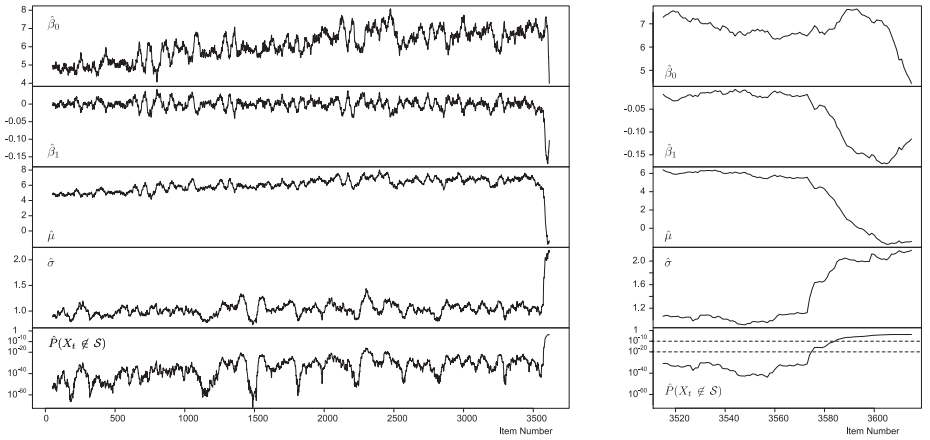

Figure 5.13: Diagnostic plots produced using our predictive method for the second day of production (depicted in Figure 5.12). The panel on the right zooms on the last 100 items produced.

which quickly exceeds 10^{-10}.

The right panel of Figure 5.13 displays the diagnostic plots for the last 100 estimates calculated before the production was halted. In a case where the stability of the machine really gets affected, the increase in $\widehat{P}(X_t \notin \mathcal{S})$ is rapid and there is only a marginal difference on the time of stoppage for any α between 10^{-12} and 10^{-4}.

The proposed method would thus have signaled despite the fact that the items were still well within the specification limits. However, the rapid changes translate into a large variance, thus a large risk which calls for an investigation.

5.6 Conclusion

Classical methods for statistical control have been designed to monitor statistical stability with limited computational resources. We propose a method that uses statistical modeling to determine the quality of the next item to be produced. The criterion that we use is similar to monitoring the instantaneous

capability of the process. In that sense, it is akin to the work of Castagliola & Vännman (2007). However, using probabilities as we do offers more flexibility, including the possibility to use covariates and to consider multivariate models. The acceptance control charts of Freund (1957) are designed for highly capable processes, but they do not offer this kind of flexibility.

The simulations and a case study illustrate that the proposed method signals timely when quality is threatened, whether the characteristics monitored are univariate or multivariate. The method is flexible, allowing a complexity in modeling that remains invisible to the user who only needs to monitor a simple univariate graph of the estimated probability of producing an item out of specifications. Furthermore, the method allows to build diagnostic plots that can provide valuable insight into the process.

The method proposed does not waive the importance of working towards ever-better stability of the production processes. It is designed to offer tools to decide if the production is of an appropriate quality rather than being geared towards detecting small (possibly insignificant) instabilities.

The review papers by Woodall & Montgomery (1999) and Woodall (2000) underline the need for new approaches in statistical quality control. The method we propose is well suited to today's technology where computerized environments allow to assess numerous items.

Woodall & Montgomery (1999) provide a list of "General Trends and Research Ideas". Our proposal offers solutions to some of these issues (the numbering used corresponds with that of Woodall & Montgomery (1999)):

(i) *"in the presence of a large amount of data [...] prevent detection of very small instabilities of no practical importance."* This is one of the main goal of our proposal.

(ii) *Using information from multi-step production.* Using our approach with models that include covariates may allow to treat such cases.

(iii) *The need for more work in detecting changes in variability.* If an appropriate model is selected, our method may not only detect changes in variability, but changes in asymmetry of the distribution too. All parameters of interest are translated in one probability, one number which summarizes consumer risk appropriately.

(v) *Challenges with multivariate control charts.* Our method can deal with multivariate distributions and even offers additional diagnostic tool to analyze the process.

At this stage, the theoretical properties of our proposal are not yet well known. However, the performances observed in simulations and in the case study combined to the flexibility and the generality of our approach are promising enough to be tested in practical situations.

Chapter 6

Monitoring a Bivariate Process using Predictive Risk

In this manuscript the method for process monitoring presented in Plante and Windfeldt (2009) is applied to a bivariate normal process. The properties of the method are explored and a case study is provided.

6.1 Introduction

In this manuscript we explore the properties of the process monitoring method introduced in Plante and Windfeldt (2009) for a bivariate normal process.

The method monitors a quality characteristic using a statistical model of the characteristic and a sliding window of observations to predict the probability that the process produce items outside the specification region. The process is stopped if this probability exceeds a pre-determined threshold. For further details on the method the reader is referred to Plante and Windfeldt (2009).

The case of a simplified sampling scheme where the predictive model is fitted on disjoint windows of observations is explored in Section 6.2. In Section 6.3 we consider the case where the model is fitted continously on a sliding window of observations. Finally a case study is provided in Section 6.4.

6.2 Properties of the Method in a Simplified Setting

Let $\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \ldots$ be a bivariate quality characteristic with specification region $\mathcal{S}$, typically $\mathcal{S} = [L_1, U_1] \times [L_2, U_2]$. Assume that the $\mathbf{X_i}$ are independent bivariate normally distributed with mean parameter $\mu = (\mu_1, \mu_2)$ and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Fixing a pre-determined threshold α, let $\mathcal{A}_\alpha$ be the acceptable operating range for the process. That is, let $\mathcal{A}_\alpha$ be the set of parameters $(\nu_1, \nu_2, \varrho, \tau_1, \tau_2)$ for which $P(\mathbf{Y} \notin \mathcal{S}) < \alpha$, where $\mathbf{Y}$ is bivariate normally distributed with mean $\nu = (\nu_1, \nu_2)$ and covariance matrix

$$T = \begin{pmatrix} \tau_1^2 & \varrho\tau_1\tau_2 \\ \varrho\tau_1\tau_2 & \tau_2^2 \end{pmatrix}. \tag{6.2.1}$$

In order to obtain a better understanding of the acceptable operating range, we decompose $\mathcal{A}_\alpha$ into smaller sets by fixing ν, i.e., we write $\mathcal{A}_\alpha = \cup_\nu \{\nu\} \times \mathcal{B}_{\alpha,\nu}$.

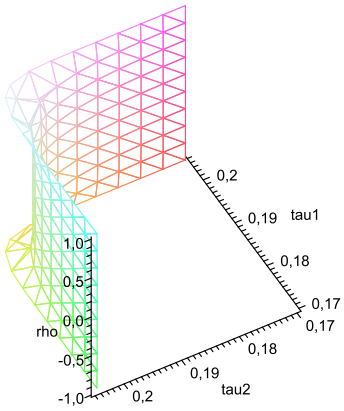

Figure 6.1: The surface that together with the $\tau_1 = 0$, $\tau_2 = 0$, $\varrho = -1$, and $\varrho_1 = 1$ planes bound the set $\mathcal{B}_{10^{-6},\,(0,0)}$.

For $\alpha = 10^{-6}$ and $\mathcal{S} = [-1, 1]^2$, an example of $\mathcal{B}_{\alpha,\,\nu}$ for $\nu = (0, 0)$ is depicted in Figure 6.1.

The control procedure is based on estimating the parameters μ and Σ by the maximum likelihood method based on a window of observations $x_1, \ldots, x_n$ of length n and signaling an alarm if $(\widehat{\mu}, \widehat{\Sigma}) \notin \mathcal{A}_\alpha$. When the model is fitted on disjoint windows of observations the estimates will be mutually independent. We are interested in the properties of this control procedure, in particular the average run length (ARL) before a signal.

Let Z be distributed like $\widehat{\mu}$, hence Z is a bivariate normally distributed variable with mean μ and covariance matrix Σ/n, and let W be distributed like $\widehat{\Sigma}$, hence W is Wishart distributed with parameters Σ and n. Then Z and W are independent and therefore

$$\begin{aligned} P((Z, W) \in \mathcal{A}_\alpha) &= \iint_{\mathbb{R}^2} \varphi_Z(\nu_1, \nu_2) \left(\iiint_{\mathcal{B}_{\alpha,\,(\nu_1,\nu_2)}} \varphi_W(T)\, d\tau_2\, d\tau_1\, d\varrho \right) d\nu_2\, d\nu_1 \\ &= E(P(W \in \mathcal{B}_{\alpha,\,Z})), \end{aligned}$$

where φ_Z and φ_W are the density functions of Z and W, respectively, and T

is the covariance matrix (6.2.1). This allows for the following expression of the average run length while under control:

$$ARL_0 = (1 - E(P(W \in \mathcal{B}_{\alpha,\, Z})))^{-1}.$$

The above mean value may be approximated by

$$E(P(W \in \mathcal{B}_{\alpha,\, Z})) \approx \frac{1}{N} \sum_{i=1}^{N} P(W \in \mathcal{B}_{\alpha,\, z_i}),$$

where are z_i are scattered appropriately and N is an integer that controls the precision of the numerical approximation. To compute this we need to determine the boundaries of $B_{\alpha,\, z_i}$, which we so far have not been able to do.

Instead we use simulations to estimate $P((\widehat{\mu}, \widehat{\Sigma}) \in \mathcal{A}_\alpha)$ an thereby the ARL. We choose $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1/6$ and $\rho = 0.5$. For $\mathcal{S} = [-1, 1]^2$ this yield a marginal capability of $C_{pk} = 2$ and $P(\mathbf{X_t} \in \mathcal{S}) = 3.95 \cdot 10^{-9}$. The size of the sliding window is set to $n = 50$ and we consider $\alpha \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-6}\}$. The length of ARL_0 makes it unsuitable for simulation.

We determine the ARL for different mean and variance values. The results are shown in Tables 6.1 and 6.2.

			ARL			
μ_1	μ_2	$P(\mathbf{X}_t \notin [-1, 1]^2)$	$\alpha = 10^{-2}$	10^{-3}	10^{-4}	10^{-6}
0	0	$3.95 \cdot 10^{-9}$	–	–	–	30.86
1/4	1/4	$6.77 \cdot 10^{-6}$	–	12500	15.94	1.14
1/2	1/2	$2.62 \cdot 10^{-3}$	25.99	1.15	1.00	1.00
1	1	$16.67 \cdot 10^{-1}$	1.00	1.00	1.00	1.00

Table 6.1: ARL for different values of μ with $\sigma_1 = \sigma_2 = 1/6$ and $\rho = 0.5$.

6.3 Illustration of the Method with Continuous Evaluations

We use simulations to explore the properties of the method when the model is fitted after each new observation. This means that we use a sliding window of

			ARL			
σ_1	σ_2	$P(\mathbf{X}_t \notin [-1,1]^2)$	$\alpha = 10^{-2}$	10^{-3}	10^{-4}	10^{-6}
1/6	1/6	$3.95 \cdot 10^{-9}$	–	–	–	31.11
1/5	1/5	$1.14 \cdot 10^{-6}$	–	100000	110.74	1.52
1/3	1/3	$5.24 \cdot 10^{-3}$	4.33	1.02	1.00	1.00
1	1	$5.02 \cdot 10^{-1}$	1.00	1.00	1.00	1.00

Table 6.2: ARL for different values of σ_1 and σ_2 with $\mu = (0, 0)$ and $\rho = 0.5$.

n observations. Using a sliding window we are able to catch unwanted changes faster. We consider different scenarios where the process losses its stability and verify that the process stops before the probability of producing an item out of specification reaches α. The number of items produced outside specifications is also investigated as well as whether the method stops the process too soon.

We consider a process assumed to be bivariate normal with initial parameters equal to $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$ and $\rho = 0.5$. We let $\mathcal{S} = [-6, 6]^2$ yielding $P(\mathbf{X_t} \notin \mathcal{S}) = 1.97 \cdot 10^{-9}$. and a marginal process capability of $C_{pk} = 2$. The maximum likelihood method is used to fit the model to the observations. All parametres are estimated in each step since fixing parameters leaves us vulnerable to changes in the fixed parameters as illustrated in Plante and Windfeldt (2009). The threshold is set to $\alpha = 10^{-4}$.

We want to detect an unwanted change as quickly as possible which suggests a small window size. On the other hand, we also want to be able to estimate the parameters in the model precisely. To be able to estimate the variance and correlation precisely the window size is set to $n = 50$.

Again, the average run length under control ARL_0 is so long that it is unsuitable for simulation. For 1000 simulated production runs of at most 10000 items, the average run length exceeded 10000 in 99.9% of the runs. In Table 6.3 are different scenarios that we have chosen to illustrate the properties of the method. There are many ways a process can loss its stability in this setting – we have chosen scenarios that we think are likely to occur in a real process. Scenarios 1a and 1b illustrate two types of shifts in the mean value. Scenarios 2a and 2b illustrate two types of variance increase. Scenarios 3a and 3b illustrate, for uncorrelated variables, a shift in the mean and variance, respectively. Finally, Scenario 4 illustrates a type of shift which occurs in the process explored in the

Scenario	μ_1	μ_2	σ_1^2	σ_2^2	ρ
1a	$i/1000$	$i/1000$	1	1	0.5
1b	0	$i/1000$	1	1	0.5
2a	0	0	$1 + i/1000$	1	0.5
2b	0	0	$1 + i/1000$	$1 + i/1000$	0.5
3a	$i/1000$	0	1	1	0
3b	0	0	$1 + i/1000$	1	0
4	0	0	$1 + i/1000$	1	$-\sigma_2/\sigma_1$

Table 6.3: Scenarios for simulations.

Scenario	1a	1b	2a	2b	3a	3b	4
	1.6	0.7	2.2	1.7	1.3	1.8	2.2

Table 6.4: Rate of items out of specifications in items per million. The values are based on 10000 runs.

case study. Scenarios 1 to 3 are illustrated graphically in Figure 6.2. Note that Scenario 4 resembels Scenario 2a.

For each scenario we simulate 10000 production runs of at most 10000 items. In Figure 6.3 is the result of the simulations. For all scenarios the method stops the process before the true probability of producing an item out of specification reaches α in almost all cases. The majority of the stoppages are close to the vertical line so false alarms are infrequent. In Table 6.4 are the number of items that did not meet the specification limits under the simulations. The number is expressed as items per million and are quite reasonable.

We conclude that for the chosen scenarios the method signals timely when the quality is threatned and that the process is allowed to continue otherwise.

6.4 A Case Study

We consider two quality characteristics (X, Y) from an industrial process that join two plastic components. The joining results in a reduction of the height of the components. The characteristic X is the height of the components after joining and Y is the height difference. As a quality measure X describe the

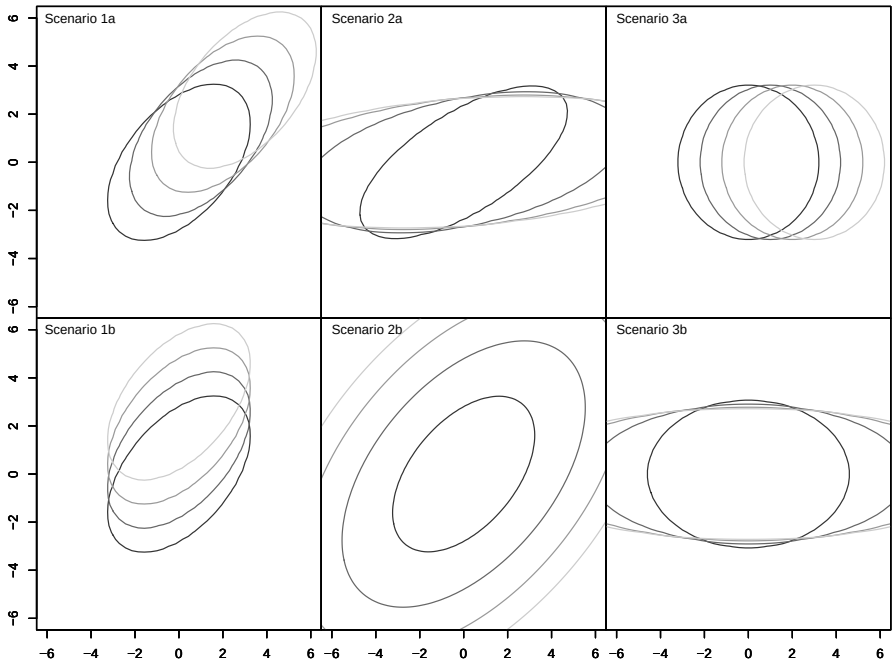

Figure 6.2: Illustration of a process lossing stability like described in the scenarios. Depicted is a fixed contour curve for the density function in each scenario. The gray colour fades as the process losses stability.

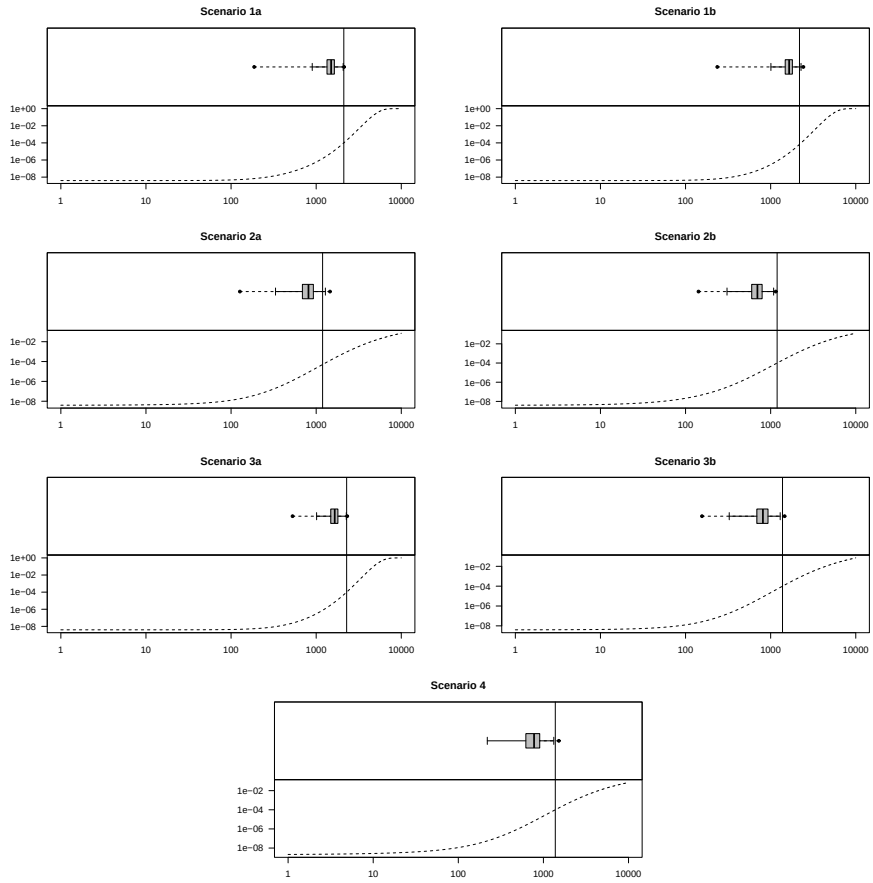

Figure 6.3: Run lengths of 10000 simulated production runs under the different scenarios from Table 6.3. Note that the log scale is used for run length. The dotted line shows the probability that the process produce an item out of specifications. The vertical line is where this probability equals $\alpha = 10^{-4}$.

geometric dimension of the joined components and Y describes the quality of the joining. The height of the components before joining Z is used to calculate Y as $Z - X$.

To get an overview of the process, we consider measurements from a span of nine production days. There is a massive amount of data, almost 150000 observations with an average of 16000 observations per day. The frequent measuring results in a lively process with e.g. monotone increases and sudden drops. Also the variation of the process seems to change over time. The process has a high capability and it is consider to be stable in a practical sense by the production engineers. The fluctuations in the process have been investigated and the causes have been determined to be either a natural part of the process or too ressource demanding to deal with. The monotone increase, for instance, is caused by the warm up of the machine during production. When the machine is stopped e.g. for maintenance or feeding, there will be a sudden drop in the level because the machine cools down. Also the slope of the increase will change after the stop.

Furthermore the components is placed in one of twelve holders on a round table when joined and measured. This results in a different level for observations coming from different holders. The levels of the holder change over time because of e.g. recalibration and cleaning which affects the scatter of the process. Furthermore, different component batches influence the level and variation of the process.

The observations from a quiet production day are depicted in Figure 6.4, where data has been rescaled to arbitrary units. For illustration purposes we have overlaid specification limits of ± 3 – the actual specification limits are wider than this. We note that there are numerous small breaks in production as well as an abrupt change at the end of the day.

To use the method we need a model for prediction. An intuitive first choice is the bivariate normal model. But because there is a difference between measurements from different holders on the round table, we need to adjust for this in the model. We choose to model holder difference as a systematic effect since we are interested in predicting the observations coming from specific holders. The holders are considered to be independent. For simplicity we suppose that the covariance structure is the same for all the holders. The model is therefore

$$(X_{ij}, Y_{ij}) \sim \mathcal{N}\left((\mu_{1j}, \mu_{2j}), \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\right), \tag{6.4.1}$$

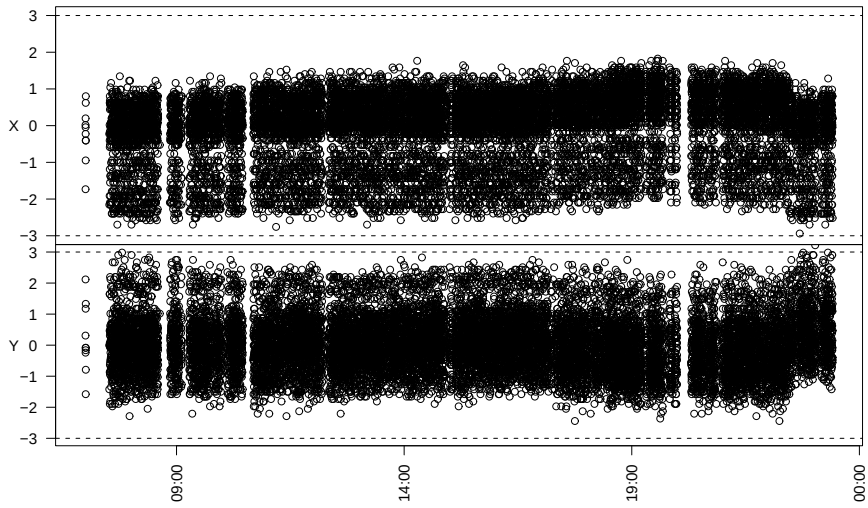

Figure 6.4: Observations on Day 1.

where $i = 1, \ldots, n$, and $j = 1, \ldots, 12$. Hence, (X_{ij}, Y_{ij}) is the i'th observation from the j'th holder.

Incorporating the monotone trends in our model will require a covariate like item number or time, but neither of these can handle the abrupt changes in the process. Handling this will require restarting the control method after each stoppage or incorporating a covariate in the model as described in Section 5 of Plante and Windfeldt (2009). Since the magnitude of the slope is very small we can – depending on the window size – make the simplifying assumption that the above model is valid across a window.

Regarding the window size we have conflicting interests. In favor of a small window size is not only the desire to avoid modeling the linear trend as described above, but also the objective of detecting unwanted changes as quickly as possible. On the other hand, we need sufficient amount of data to estimate the parameters precisely, which favors a bigger window size. A window size of $n = 600$ observations will base the mean estimates in the model on the same number of observations as in the simulations in Section 6.3. Since the covariance

structure is the same for all holders an even smaller window size will probably be sufficient.

The assumptions of the model is verified by picking a number of data windows from Day 1. For each set of observations the standardized residuals is plotted in a qq-plot and the covariance and correlation is compared across holders. Furthermore the absence of linear trend is confirmed.

The method uses a sliding window with continuous evalution after each item to detect unwanted changes as quickly as possible. However for illustration purposes we have chosen to use disjoint data windows below. This is equivalent to displaying every 50th evaluation with continuous evaluations.

This illustration does not focus so much on the specific threshold but more on the changes in $P(\mathbf{X_t} \notin \mathcal{S})$ compared to the changes in the process.

In Figure 6.5 is depicted $\widehat{P}(\mathbf{X_t} \notin \mathcal{S})$ for Day 1 together with the corresponding diagnostic charts of the estimates with 95% confidence limits. Note that it is the time of the first observation in the data window that determines the points location on the time scale.

The variance estimates are quite stable and $\widehat{\sigma_1}$ and $\widehat{\sigma_2}$ are around 0.225 and 0.325, respectively. Also the correlation is reasonable stable and is around -0.45. It is clear from the mean estimates that the holders have different levels as expected. The mean level of the holders drift during the day which is reflected in $\widehat{P}(\mathbf{X_t} \notin \mathcal{S})$. The shift in the observations at the end of the day is reflected in all the parameter estimates. When the data window contains observations from before and after the shift, the variance estimates will naturally increase. This, together with the increase in mean for Y, results in a swift increase in the estimate of $P(\mathbf{X_t} \notin \mathcal{S})$. At the next point on the charts, the data window only contains observations from after the shift and the variances are back to normal. The estimate of $P(\mathbf{X_t} \notin \mathcal{S})$ is still increasing though because the mean of Y is getting close to the specification limits.

We now go on to investigate how the chart behaves with two types of changes in the process. The first change is a change in the level which we have seen a small example of at the end of Day 1. The second example is a change in variance.

In Figure 6.6 is depicted the process on another day, Day 2. The scatter of the process over the specification region is less then on Day 1. There is an abrupt

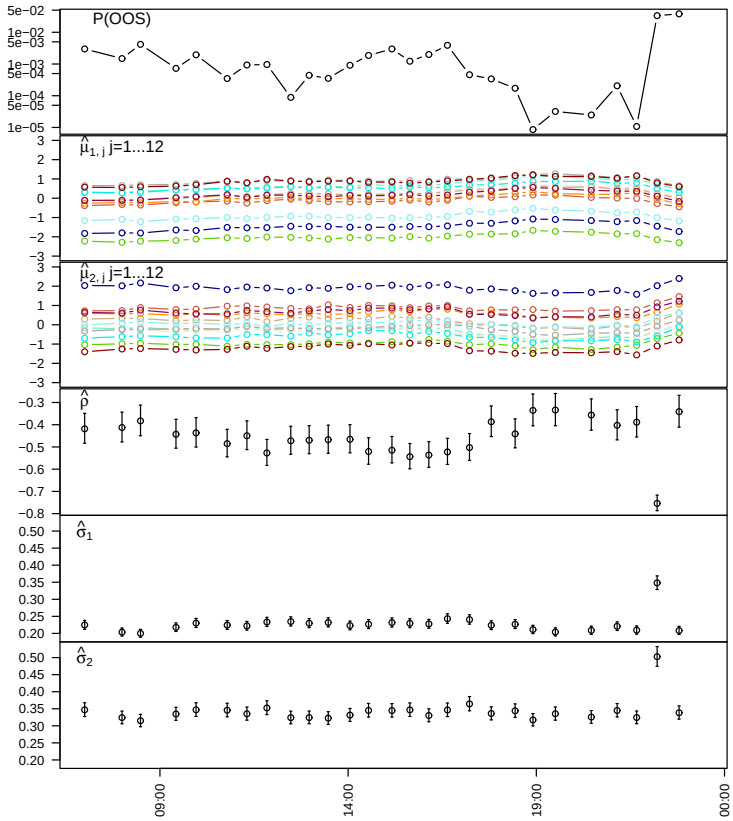

Figure 6.5: OOS chart and corresponding diagnostic chart for Day 1 based on the bivariate normal model. The window size is 50 and there is no overlap between windows. Note that we have used a log scale for the OOS chart to magnify the variation.

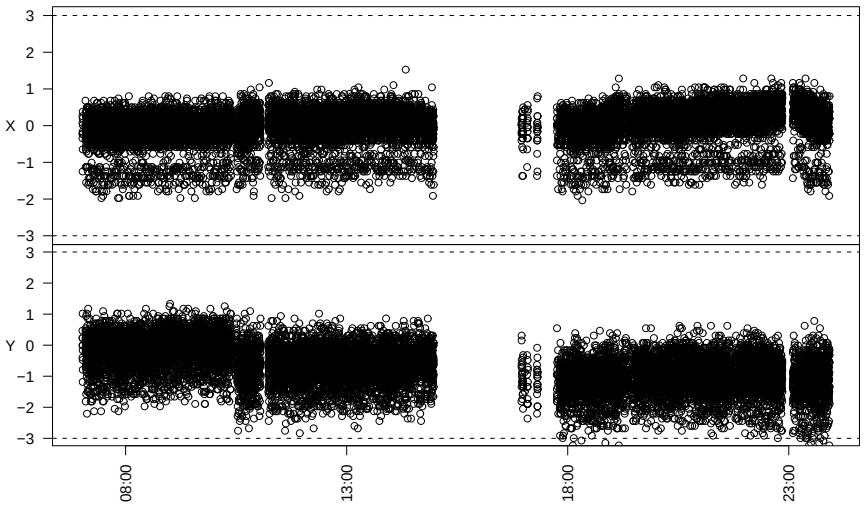

Figure 6.6: Observations on Day 2.

change mainly in Y in the late morning and after the long break in the afternoon Y seems to drop again. After a short break in production Y drops even further. Also X experiences some fluctuation during the day, but remains in the middle of the specification interval.

In Figure 6.7 is depicted the values of $\widehat{P}(\mathbf{X_t} \notin \mathcal{S})$ for Day 2 together with the corresponding diagnostic charts. The process is also depicted to easier follow the simultaneous movement.

The sudden drop in Y in the late morning is picked up by the estimate of $P(\mathbf{X_t} \notin \mathcal{S})$ which rises steeply. The decrease of Y during the rest of the day is also reflected by the estimate of $P(\mathbf{X_t} \notin \mathcal{S})$. The diagnostic charts show that the first sudden drop is caused by a drop in the levels for all holders while the worsening during the rest of the day is mainly caused by observations from one holder which are getting dangerously close to the specification limits. To understand how Y can decrease while X remains constant we return to the definition of Y. Because $Y = Z - X$ and X remain constant the increase in Y most be caused by an increase in Z. This also means that the joining of the

components is compressing them to such an extend that the increase in Z has little impact on X. Scenario 1b in Section 6.3 illustrates this type of change in a process. The diagnostic charts also show that the holder difference is much smaller than on Day 1 which explains the lower scatter in the process.

In Figure 6.8 is depicted the observations from a third day, Day 3. In the late morning Y is changing in a way that result in more observations in the upper part of the specification interval. The estimates of $P(\mathbf{X_t} \notin \mathcal{S})$ and the corresponding diagnostic charts are depicted in Figure 6.9.

The estimate of $P(\mathbf{X_t} \notin \mathcal{S})$ picks up on this change and the diagnostic charts show that there is an increase in the variance of Y. The variance of X is also increasing but not to the same extend. The increase in the variance of Y causes the correlation to decrease. The explanation for this can be found in the definition of Y and the small correlation between X and Z. Hence, the correlation between X and Y is approximately $-\sigma_1/\sigma_2$. We note that Scenario 4 in Section 6.3 illustrates the change experienced here.

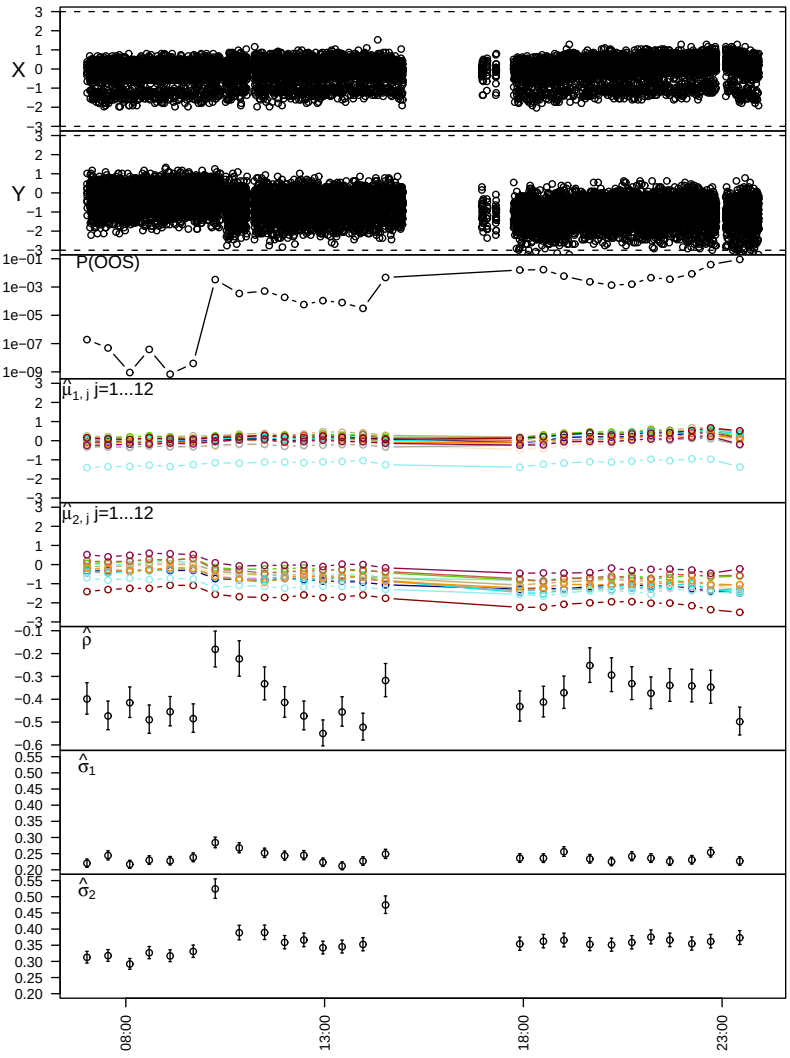

Figure 6.7: OOS chart and corresponding diagnostic charts for Day 2 based on the bivariate normal model. The window size is 50 and there is no overlap between windows. Note that we have used a log scale for the OOS chart to magnify the variation.

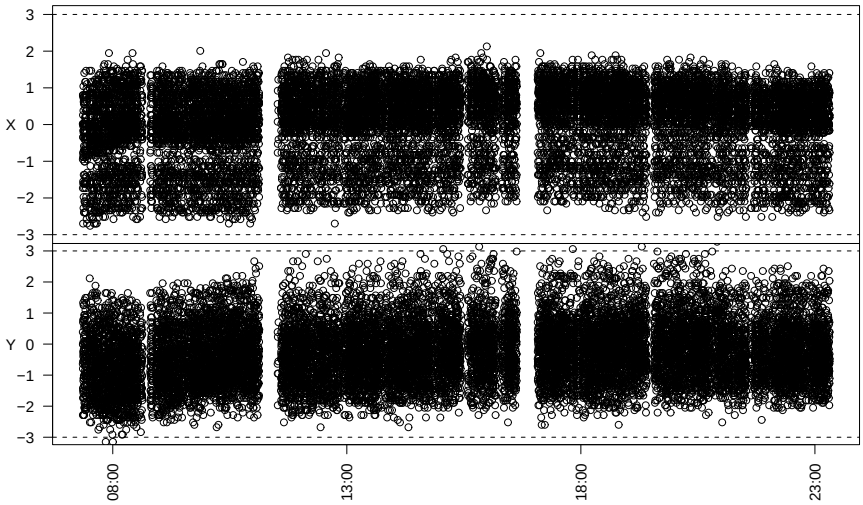

Figure 6.8: Observations on Day 3.

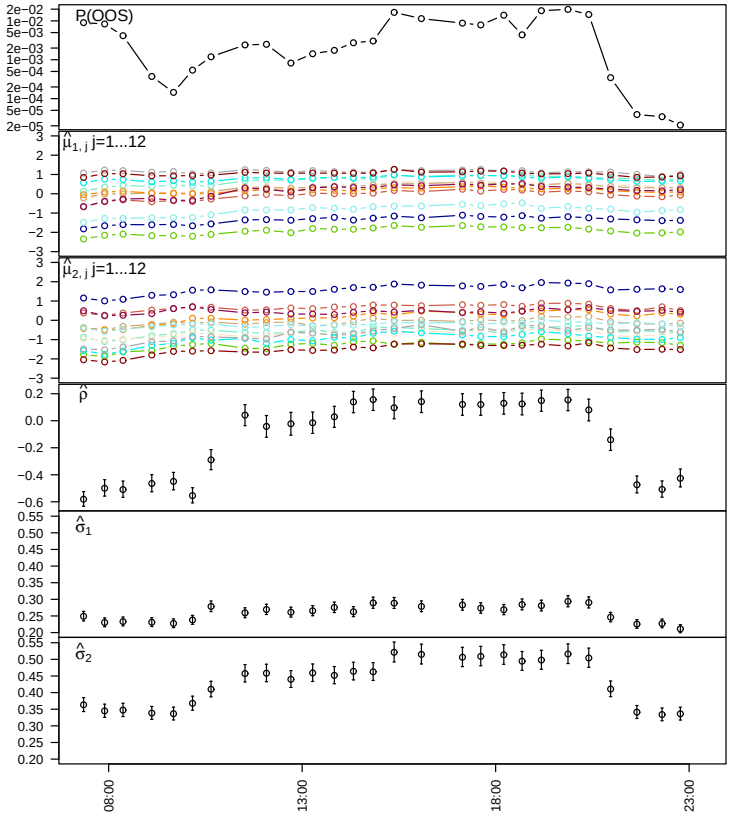

Figure 6.9: OOS chart and corresponding diagnostic charts for Day 3 based on the bivariate normal model. The window size is 50 and there is no overlap between windows.

IV

Missing Values

Chapter 7

Assessing the Impact of Missing Values on Quality Measures of an Industrial Process

Problem: In the assembly of a medical device, a measuring process was discovered to introduce an unwanted transformation. This could seriously impact the estimates of the descriptive measures, like the mean and standard deviation, that are used to make important decisions about quality.
Approach: An investigation was carried out to determine the cause of the problem. The main cause was identified, but it proved very challenging to determine the exact transformation. Instead the observed data was examined to determine likely preimages for the observed data. Based on these preimages the EM algorithm was used to find the maximum likelihood estimates for the mean and standard deviation of the measuring process.
Results: It was concluded that the transformation had no significant influence on the estimates of the mean and standard deviation of the process. Furthermore, guidelines was given along with software so the production engineers can handle similar problems in the future.

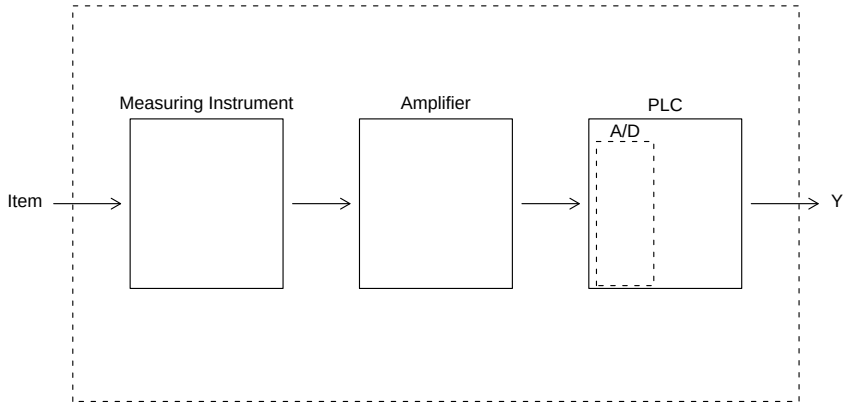

Figure 7.1: Diagram of the measuring process.

7.1 Process Description

We consider a height measurement that takes place during assembly of a medical device. The outcome of the measurement is used to monitor the quality of a subprocess of the assembly process. This type of measurement is performed several places in the assembly process.

The item about to be measured is placed in a holder on a round table. The round table is equipped with a total of 12 holders. An instrument measures the height of the item in the holder. This results in an analogue signal that is transmitted to an amplifier which sends the amplified signal to a programmable logic controller (PLC). An analogue to digital converter (AD converter) in the PLC converts the analogue signal to a digital signal in the form of an integer. Using a conversion line, the corresponding height (in millimeters) is computed by the PLC. This is the outcome of the measurement. A diagram of the measuring process is shown in Figure 7.1.

To ensure the quality of the assembly process the height measurement is moni-

tored with an $\bar{X}$-chart and an R-chart. A subgroup consists of a measurement from each holder on the round table. The limits of the $\bar{X}$-chart are based on the long term standard deviation divided with $\sqrt{12}$, i.e. the limits of a Levey-Jennings chart adjusted for the size of the subgroup. The limits of the $\bar{X}$-chart are therefore given by

$$LCL = \mu - 3\frac{\sigma}{\sqrt{12}} \quad \text{and} \quad UCL = \mu + 3\frac{\sigma}{\sqrt{12}},$$

where μ and σ are the mean and long term standard deviation of the process, respectively. Furthermore two capability indices are estimated on a daily basis to monitor the process performance. The capability indices used are

$$C_p = \frac{USL - LSL}{6\sigma} \quad \text{and} \quad C_{pk} = \min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right).$$

7.2 Problem Description

During an investigation of how to improve the control system of the assembly process, the output Y of the height measurement from one holder was examined. Figure 7.2 shows the measurements in question. It was noted that even though the result of the height measurement had a resolution of three decimals, not all possible outcomes were observed. It was found that all the holders had the same problem, however, the values missing was not the same. This fact made it impossible to discover the problem when viewing the measurements from all holders simultaneously.

The findings was discussed among the production engineers who concluded that it was not the measured items that caused the observed pattern – the items could be considered to be normally distributed. This meant that the measurements were somehow transformed during the measuring process. The estimates of the mean and standard deviation of the height measurement are used for statistical process control to ensure the quality of the product. Now the question was, what was the likely cause of the transformation and was the estimated mean and standard deviation seriously affected by this, and, if so, what could be done to handle this problem. Furthermore, because it was considered to monitor each holder separately, it was also questioned how the unwanted transformation would effect this new control strategy.

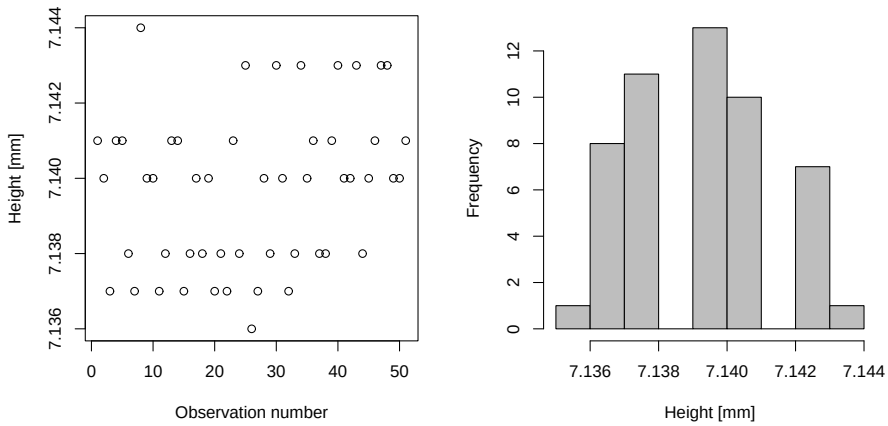

Figure 7.2: The plot on the left shows the height measurements from one holder where values are missing. On the right is a histogram of the same measurements making the missing values even more obvious.

7.3 Data Collection

We wish to determine the maximum likelihood estimate (MLE) of the mean and standard deviation of the process. Then we consider the difference between these estimates and the empirical estimates, and the difference this induces on the quality measures that are used.

The outcome of the measuring process is collected on-line and stored in a database. The measurement value is stored together with characteristics for the measurement, e.g. holder number. The sample we use for estimation is taken from a stable part of the daily production.

The impact of the transformation is related to the size of the standard deviation of the process. The smaller the standard deviation, the greater the impact, so by choosing a small sample we would see a bigger impact than with a larger sample. On the other hand, the sample should be big enough to estimate the parameters in the model precisely, also in the model where each holder is con-

sidered separately. Furthermore, since there are 12 holders it is reasonable to choose a sample size as a multiple of 12. Taking all this into consideration we have chosen a sample consisting of 600 observations. The observations can be found in Appendix B.

7.4 Analysis and Interpretation

7.4.1 Model

In order to analyze the data, we formulate a statistical model. Let $Y = (Y_1, \ldots, Y_n)$ denote the observed output from the measuring process, and let $X = (X_1, \ldots, X_n)$ denote the corresponding variable which would have been observed without the transformation. Assume that there exists a transformation $t : \mathbb{R} \to \mathbb{R}$ such that $t(X_i) = Y_i$. Assuming that the untransformed variables X_i are independent normally distributed with mean μ and variance σ^2, the likelihood function for X is given by

$$L_X(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \; e^{\,-\frac{1}{2\sigma^2}(X_i-\mu)^2},$$

and the log likelihood function is given by

$$l_X(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2. \tag{7.4.1}$$

If we had observed X, we would find the MLE by maximizing the log likelihood function to obtain

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \widehat{\mu})^2.$$

In our case, however, we only have the observed Y. The likelihood function for Y is given by

$$\begin{aligned} L_Y(\mu, \sigma^2) &= \prod_{i=1}^{n} P\left(X_i \in t^{-1}(Y_i)\right) \\ &= \prod_{i=1}^{n} \int_{t^{-1}(Y_i)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\,-\frac{1}{2\sigma^2}(x-\mu)^2} dx. \end{aligned}$$

It follows that the log likelihood function is given by

$$l_Y(\mu, \sigma^2) = \sum_{i=1}^{n} \log \left(\int_{t^{-1}(Y_i)} \frac{1}{\sqrt{2\pi\sigma^2}} \; e^{\; -\frac{1}{2\sigma^2}(x-\mu)^2} dx \right).$$

To find the MLE for Y we need to maximize l_Y, but in order to do so we need to determine the preimages $t^{-1}(Y_i)$ of the observations.

7.4.2 Identification of Preimages

We began to determine the transformation t by studying the measuring process in detail. This approach proved to be a major challenge and required help from many different specialists. In the process we obtained a good understanding of the cause of the problem.

We began by studying the machine documentation to determine which instruments were used in the measuring process. Here we found the type of the measuring instrument as well as the instruments used for signal processing. We confirmed the findings by locating the actual instruments on the machine. The instruments come in different makes and the production engineers helped to determine the specific types. A diagram of the instruments in the measuring process is shown in Figure 7.1.

After going through the manuals of the instruments and consulting engineers specialized in measuring processes, we decided to view the process as consisting of two parts. The first is the actual measuring consisting of the measuring instrument and analogue signal processing in the amplifier. The second part is the analogue to digital conversion and the subsequent data processing taking place in the PLC. Given the systematic nature of the problem the attention was focused on the second part where the problem most likely originates.

We break the second part in two. First the conversion from analogue to digital signal and second, the conversion of this digital signal to the observed value.

The analogue to digital conversion is done by an AD converter. To understand this better we consulted the manuals and the programmer that programmed the PLC. We learned that the output from the amplifier is in the range from -10V to 10V. Furthermore we learned that the AD converter converts this output to an integer between -27648 and 27648. The ideal transfer function

i	$int_{low,i}$	$int_{high,i}$	i	$int_{low,i}$	$int_{high,i}$
1	556	3654	7	574	3672
2	570	3670	8	592	3694
3	580	3680	9	586	3690
4	558	3658	10	610	3710
5	552	3654	11	598	3698
6	584	3684	12	594	3694

Table 7.1: The integer values corresponding to the gauge blocks for all holders.

ad can be described as $ad(x) = \text{trunc}(2764.8 \cdot x)$, where trunc is the function that truncates the argument to the nearest integer towards zero. Note that the digital representation of the analogue value is restricted to be an integer. It is the programmers opinion that the problem we are experiencing is caused by the resolution of the AD converter and possibly the subsequent calculation.

The observed value (in millimeters) is calculated from the integer value with the function

$$c(x) = \text{trunc}\left((x - int_{low,i}) \cdot \frac{h_{high} - h_{low}}{int_{high,i} - int_{low,i}} - h_{low}\right) \cdot 10^{-3},$$

where $h_{low} = 6501\ \mu m$ and $h_{high} = 8752\ \mu m$ are the heights of the two gauge blocks that are used for calibration and $int_{low,i}$ and $int_{high,i}$, where $i = 1, \ldots, 12$ are the corresponding integer values for each holder. The machine is calibrated when needed e.g. if a holder is replaced. When the machine is recalibrated the values of $int_{low,i}$ and $int_{high,i}$ typically changes. The values of $int_{low,i}$ and $int_{high,i}$ corresponding to the sample considered in this paper are given in Table 7.1.

This means that we can describe the second part of the measuring process with the composite function $c \circ ad$. Unfortunately, using this function alone to describe the second part of the measuring process does not result in the missing values we observe. A more in depth investigation to find a more accurate description of the transfer function is therefore required. However, from a resource perspective, it was decided to try a different approach before possibly continuing with this.

We began by looking at the values present in the observed data. These are

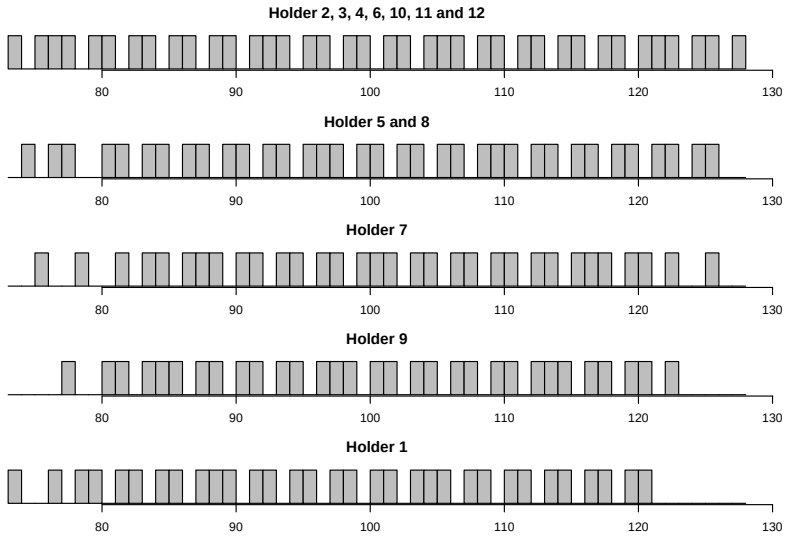

Figure 7.3: Values present for the 12 holders.

depicted in Figure 7.3 where they have been transformed using the function $f(x) = x \cdot 10^3 - 7000$. The holders that have the same values present are depicted together. We focused on the middle part of the interval because the pattern of missing and present values may be incomplete at the edges due to too few observations. The values present at the edges should of course match the pattern deducted from the middle part, but a missing value in the outer edges is not necessarily caused by the transformation. We can shift the observed values so the holders have exactly the same values present except holder 1 which has a slightly different pattern. The shift needed for each holder is given in Table 7.2. We depict all the present values from holders 2 to 12 together to get a clearer understanding of the pattern, this is done in Figure 7.4. We have removed the values since they do not have any meaning after the translation.

The preimages of the transformation is a grouping of the exact values and the observed value is the representation of the exact values in the group. It is intuitive to restrict ourselves to think of the preimages as disjoint intervals with the same length. Since some values are missing we would expect that the size

Holder	Shift
$2, 3, 4, 6, 10, 11$ and 12	0
5 and 8	-4
7	$+5$
9	$+8$

Table 7.2: The shift needed to align the present values for holders 2 to 12.

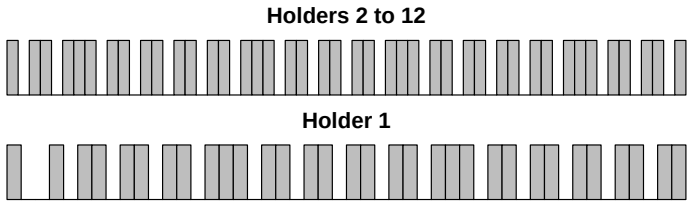

Figure 7.4: Pattern of missing and present values for holders 2 to 12 and holder 1, respectively.

of interval of exact values for a given observed value is larger than one. An interval of length 1.45 gives us the pattern experienced in the data for holder 2 to 12. This value is not unique since values lying in a small interval around 1.45 also produce the same pattern. For holder 1 an interval length 1.454 gives the observed pattern. The preimages of the holders 1 and 2 are can be found in Appendix B and are depicted in Figure 7.5. The preimages for the rest of the holders are found by shifting the intervals for holder 2 appropriately, these can also be found in Appendix B.

7.4.3 Estimation

Since this is essentially is a missing data problem, it is natural to use the EM algorithm to maximize the log likelihood function, see Dempster et al. (1977). An alternative would be to maximize the log likelihood function directly using a numerical maximization routine, for instance the Newton-Raphson algorithm. This would probably give a quicker convergence, but also a more unstable algorithm that is harder to implement in basic software. A good introduction to

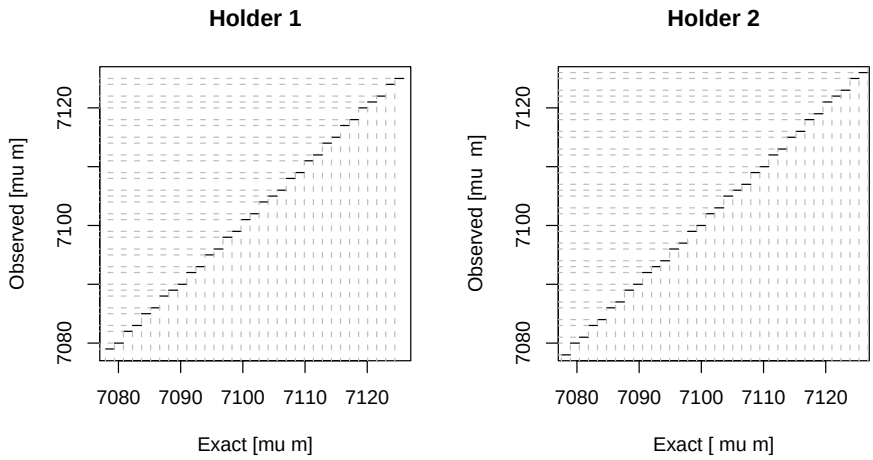

Figure 7.5: The preimages of the observation for holders 1 and 2.

the EM algorithm can be found in McLachlan and Krishnan (1997).

Using the empirical estimates as staring point, we look at the EM algorithm in the k'th iteration. Assume we are given a set of parameters $\theta_k = (\mu_k, \sigma_k^2)$ and we want to determine $E_{\theta_k}\left(l_X\left(\theta\right)|X_i \in I_i\right)$, where $I_i = [L_i, U_i[$ is the preimage of $t^{-1}(Y_i)$. From Equation (7.4.1), we get

$$\begin{aligned} &E_{\theta_k}\left(l_X\left(\theta\right)|X_i \in I_i\right) \\ &\qquad = E_{\theta_k}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\middle| X_i \in I_i\right). \end{aligned}$$

We have

$$E_{\theta_k}\left((X_i-\mu)^2|X_i \in I_i\right) = E_{\theta_k}\left((X_i-\mu)|X_i \in I_i\right)^2 + V_{\theta_k}\left((X_i-\mu)|X_i \in I_i\right).$$

For short, let

$$\begin{aligned} [X_i]_k &= E_{\theta_k}\left(X_i|X_i \in I_i\right), \\ [X_i^2]_k &= E_{\theta_k}\left(X_i^2|X_i \in I_i\right). \end{aligned}$$

We may then write

$$E_{\theta_k}\left((X_i-\mu)^2|X_i\in I_i\right)=\left([X_i]_k-\mu\right)^2+[X_i^2]_k-[X_i]_k^2,$$

and so

$$\begin{aligned}&E_{\theta_k}\left(l_X\left(\theta\right)|Y\right)\\&\quad=-\frac{n}{2}\log(2\pi)-\frac{n}{2}\log(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\left([X_i]_k-\mu\right)^2+[X_i^2]_k-[X_i]_k^2\right).\end{aligned}$$

We now want to maximize this function with respect to μ and σ^2. We note that the function is equivalent to the log likelihood function for X, except for the extra term $[X_i^2]_k-[X_i]_k^2$. This term is invariant when maximizing with respect to μ, but has to be taken into account when maximizing with respect to σ^2. The MLE for the expected log likelihood is therefore given by

$$\widehat{\mu}_*=\frac{1}{n}\sum_{i=1}^{n}[X_i]_k\quad\text{and}\quad\widehat{\sigma^2}_*=\frac{1}{n}\sum_{i=1}^{n}\left(\left([X_i]_k-\overline{[X_i]}_k\right)^2+[X_i^2]_k-[X_i]_k^2\right).$$

This gives us $\mu_{k+1}=\widehat{\mu}_*$ and $\sigma_{k+1}^2=\widehat{\sigma^2}_*$ for the next step in the algorithm.

The EM algorithm does not automatically provide us with the variance of the estimates. But if we let

$$s_i(\widehat{\theta})=\frac{\partial\log P_\theta(X_i\in I_i)}{\partial\theta}\Big|_{\theta=\widehat{\theta}},$$

then an approximation to the observed information is given by

$$I(\widehat{\theta},Y)\approx\sum_{i=1}^{n}s_i(\widehat{\theta})s_i(\widehat{\theta})^T,$$

Hence, an approximation to the covariance matrix is given by $I(\widehat{\theta},Y)^{-1}$. A detailed argument is given in Sections 4.3 and 4.4 of McLachlan and Krishnan (1997) and Jones and Mclachlan (1992).

7.4.4 Implementation

The calculation of the expected mean of X_i and X_i^2 involves evaluating an integral. We give formulas that only depend on the density and the cumulative

distribution function of a normal distribution, which comes as standard in many non-statistical software packages, e.g. Excel. We have

$$E_{(\mu,\sigma^2)}(X_i|X_i \in [L_i, U_i]) = \mu - \sigma^2 \frac{\varphi_{\mu,\sigma^2}(U_i) - \varphi_{\mu,\sigma^2}(L_i)}{\Phi_{\mu,\sigma^2}(U_i) - \Phi_{\mu,\sigma^2}(L_i)},$$

and

$$E_{(\mu,\sigma^2)}(X_i^2|X_i \in [L_i, U_i]) = \mu^2 + \sigma^2 - \sigma^2 \frac{(U_i+\mu)\varphi_{\mu,\sigma^2}(U_i) - (L_i+\mu)\varphi_{\mu,\sigma^2}(L_i)}{\Phi_{\mu,\sigma^2}(U_i) - \Phi_{\mu,\sigma^2}(L_i)},$$

where φ_{μ,σ^2} and Φ_{μ,σ^2} are the density and cumulative distribution function for a normal distribution with mean μ and variance σ^2, respectively.

Based on these formulas we can derive another set of formulas for the partial derivatives of the log likelihood function to calculate the variance of the ML estimate. These are given by

$$\begin{aligned}
\frac{\partial \log P_{(\mu,\sigma^2)}(X_i \in I_i)}{\partial \mu} &= -\frac{\varphi_{\mu,\sigma^2}(U_i) - \varphi_{\mu,\sigma^2}(L_i)}{\Phi_{\mu,\sigma^2}(U_i) - \Phi_{\mu,\sigma^2}(L_i)}, \\
\frac{\partial \log P_{(\mu,\sigma^2)}(X_i \in I_i)}{\partial \sigma^2} &= -\frac{1}{2\sigma^2} \frac{(U_i-\mu)\varphi_{\mu,\sigma^2}(U_i) - (L_i-\mu)\varphi_{\mu,\sigma^2}(L_i)}{\Phi_{\mu,\sigma^2}(U_i) - \Phi_{\mu,\sigma^2}(L_i)}.
\end{aligned}$$

We have implemented the EM algorithm in the statistical software `R` and in Visual Basic as a macro for use in Excel. The code can be found in Appendix A.

7.4.5 Results and Interpretation

Having determined likely preimages for the observations we can calculate the log likelihood function. This is illustrated in Figure 7.6.

We use the EM algorithm to maximize the log likelihood function, using the empirical estimates as starting point. In Table 7.3 is given the mean and standard deviation estimates and the corresponding value of the log likelihood function after each iteration k.

We see that the EM algorithm converges quickly. The estimates are given in Table 7.4. Regarding the new control strategy where each holder is be considered separately, we find the estimates for each holder by using the EM algorithm on

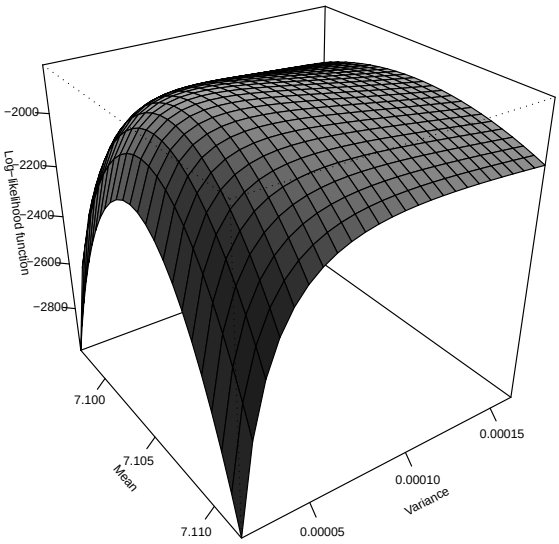

Figure 7.6: Log likelihood function for the sample.

k	μ_k	σ_k	$l(\mu_k, \sigma_k)$	$l(\mu_k, \sigma_k) - l(\mu_{k-1}, \sigma_{k-1})$
0	7.104853	0.00735766	-1826.974	
1	7.104644	0.00735802	-1826.731	0.24
2	7.104644	0.007358022	-1826.731	$2.53 \cdot 10^{-6}$
3	7.104644	0.007358022	-1826.731	$2.66 \cdot 10^{-11}$

Table 7.3: Convergence of the EM algorithm for all holders.

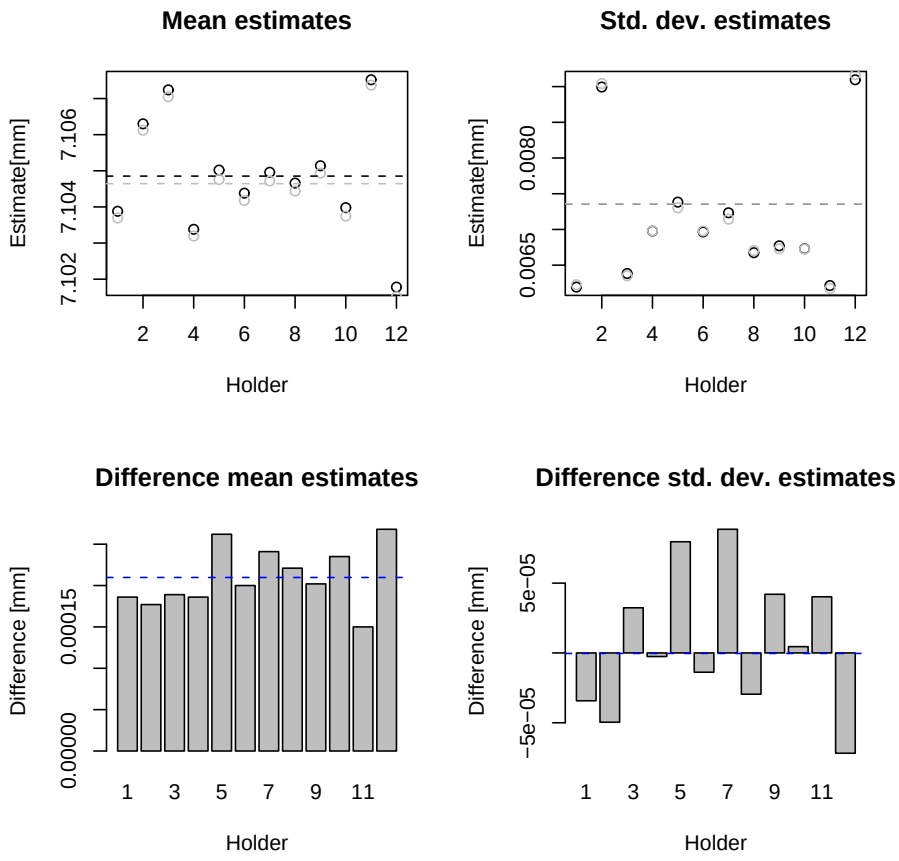

Figure 7.7: Empirical and ML estimates of the mean and standard deviation of the observations. In the two upper figures the black color represents the empirical estimates and the gray the ML estimates. The dotted lines are the estimates based on all observations. In the two lower figures the dotted lines represent the difference of the estimates based on all observations.

Holder	Mean			Std. dev.		
	Emp.	MLE (10^{-4})	Diff. (10^{-4})	Emp. (10^{-3})	MLE (10^{-3}) (10^{-6})	Diff. (10^{-3})
All	7.10485	7.10464 (3.1)	2.1	7.3577	7.3580 (3.0)	−0.0004
1	7.10388	7.10369 (9.1)	1.9	6.1924	6.2266 (7.6)	−0.0343
2	7.10630	7.10612 (7.4)	1.8	8.9939	9.0435 (7.0)	−0.0496
3	7.10724	7.10705 (5.7)	1.9	6.3830	6.3506 (5.4)	0.0323
4	7.10338	7.10319 (5.0)	1.9	6.9768	6.9794 (4.6)	−0.0026
5	7.10502	7.10476 (4.5)	2.6	7.3824	7.3027 (4.1)	0.0796
6	7.10438	7.10418 (4.1)	2.0	6.9653	6.9791 (3.8)	−0.0138
7	7.10496	7.10472 (3.8)	2.4	7.2331	7.1446 (3.5)	0.0886
8	7.10466	7.10444 (3.6)	2.2	6.6742	6.7037 (3.2)	−0.0295
9	7.10514	7.10494 (3.3)	2.0	6.7706	6.7285 (3.0)	0.0420
10	7.10398	7.10374 (3.2)	2.4	6.7305	6.7260 (2.8)	0.0045
11	7.10752	7.10737 (3.0)	1.5	6.2137	6.1734 (2.7)	0.0402
12	7.10178	7.10151 (2.9)	2.7	9.0957	9.1675 (2.6)	−0.0718

Table 7.4: Empirical and ML estimates of the mean and standard deviation of the observations. The number in parenthesis after each ML estimate is the standard deviation of the estimate.

the subset of observations from each holder. These estimates can also be found in Table 7.4. The results are displayed graphically in Figure 7.7.

As we can see the difference in the two mean estimates is small. The empirical mean is generally slightly over estimated since the observed values tend to lie in the upper range of the preimage interval. The difference between the mean estimates, when considering each holder separately, are roughly the same as the difference in the simple model.

The difference in the standard deviation in the simple model is very small. As we would expect the impact is greater, when viewing each holder separately, because the variability is smaller relatively to the length of the preimage intervals.

To determine the impact on the quality measures we calculate the values of the upper and lower control limits of the $\bar{X}$–chart as well as the capability indices C_p and C_{pk} for both estimates. The values of the control limits are inserted in Table 7.5 and the capability indices are inserted in Table 7.6. The differences are depicted in Figure 7.8.

The small difference of the mean and standard deviation estimates and the high capability of the process makes the difference in the capability indices

Holder	LCL Emp.	LCL MLE	LCL Diff. (10^{-4})	UCL Emp.	UCL MLE	UCL Diff. (10^{-4})
All	7.09848	7.09827	2.1	7.11123	7.11102	2.1
1	7.09852	7.09830	2.2	7.10924	7.10909	1.6
2	7.09851	7.09829	2.2	7.11409	7.11395	1.3
3	7.10171	7.10155	1.6	7.11277	7.11255	2.2
4	7.09734	7.09715	1.9	7.10942	7.10924	1.8
5	7.09863	7.09843	1.9	7.11141	7.11108	3.3
6	7.09835	7.09814	2.1	7.11041	7.11022	1.9
7	7.09870	7.09853	1.6	7.11122	7.11091	3.2
8	7.09888	7.09863	2.5	7.11044	7.11024	2.0
9	7.09928	7.09911	1.7	7.11100	7.11077	2.4
10	7.09815	7.09792	2.3	7.10981	7.10957	2.4
11	7.10214	7.10202	1.2	7.11290	7.11272	1.8
12	7.09390	7.09357	3.3	7.10966	7.10945	2.1

Table 7.5: Upper and lower control limits of the $\bar{X}$-chart based on the empirical and ML estimates.

Holder	C_p Emp.	C_p MLE	C_p Diff.	C_{pk} Emp.	C_{pk} MLE	C_{pk} Diff.
All	13.455	13.455	0.001	11.184	11.173	0.010
1	15.987	15.899	0.088	13.288	13.204	0.084
2	11.007	10.947	0.060	9.149	9.091	0.058
3	15.510	15.589	−0.079	12.891	12.946	−0.055
4	14.190	14.185	0.005	11.794	11.780	0.014
5	13.410	13.557	−0.146	11.146	11.258	−0.112
6	14.213	14.185	0.028	11.813	11.780	0.033
7	13.687	13.857	−0.170	11.376	11.507	−0.131
8	14.833	14.768	0.065	12.329	12.264	0.065
9	14.622	14.713	−0.091	12.153	12.219	−0.065
10	14.709	14.719	−0.010	12.226	12.223	0.002
11	15.933	16.036	−0.104	13.243	13.317	−0.075
12	10.884	10.799	0.085	9.047	8.968	0.078

Table 7.6: C_p and C_{pk} based on the empirical and ML estimates.

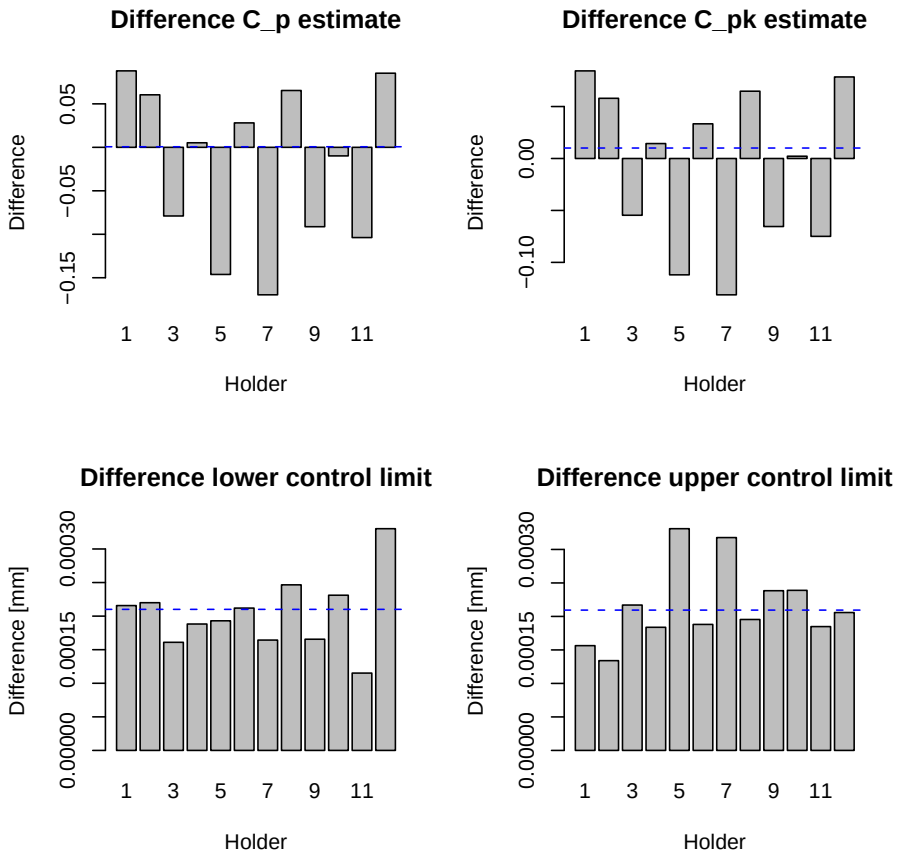

Figure 7.8: Difference between quality measures. The dotted line is the difference between the estimates that are based on all observations.

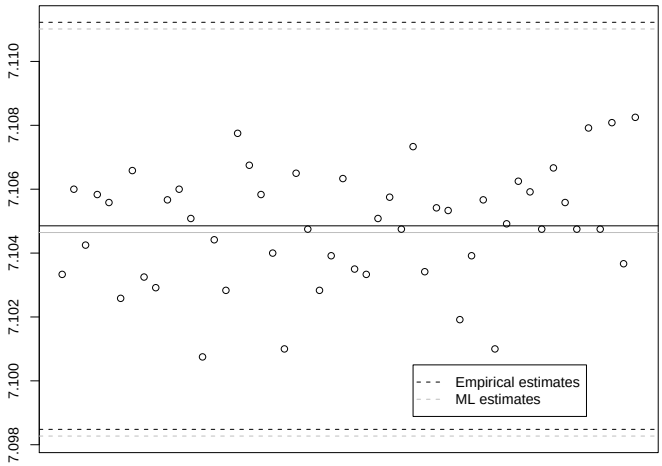

Figure 7.9: An $\bar{X}$-chart with two sets of limits and two mean values depicted based on the estimates calculated from all observations. The black color represents the empirical estimates and the gray is the ML estimates. The points depicted are calculated from the sample.

insignificant for the simple model. Due to the high capability of the process we still conclude that the difference – when considering each holder separately – is of no practical importance.

The difference in the values of the control limits are roughly the same when viewing each holder separately as in the simple model. In Figure 7.9 is depicted an $\bar{X}$–chart of the measurements with both set of limits drawn based on the simple model. We conclude that the impact on the control limits is not practically significant for this process.

7.5 Conclusion and Recommendations

As we have seen there is no practical difference between the empirical estimates and the ML estimates found by the EM algorithm in the context they are used today. Hence, we conclude that the current process controls are acceptable, despite the problem with the measurement system.

We can consider whether we would have reached a different conclusion had we obtained the actual preimages. The answer is most likely not. The interval lengths of the exact preimages will not necessarily be identical, as we have assumed. They could, for instance, be centered differently, but this will not significantly affect the estimates.

We could determine the exact transformation, but this will require that we disassemble the machine and halt the production during this time. Since the value of this is most likely very limited, we will not proceed further in the pursuit of the exact transformation of the measurement process.

The problem and the subsequent investigation, have made us aware of the risk of introduction of unwanted transformation in the signal processing of a measuring process. This type of signal processing is very common and therefore a method has been provided along with software to handle similar problems in the future. The method relies on finding likely preimages of the transformation for the observations and we have shown two different approaches to finding these preimages. Further we have seen the importance of detailed documentation of the measuring process, since it is very resource demanding to reconstruct this later.

The method provided in this paper can also be used to handle the special case of normal number rounding. In this case, the preimages are given by $[x - 5 \cdot 10^{-d}, x + 5 \cdot 10^{-d}[$, where x is the observed value and d is the number of digits rounded to. Another very simple way to handle this problem is to use Sheppard's correction as described in Stuart and Ord (1998). Using Sheppard's correction, the mean value will remain the same while the variance will be corrected with $-h^2/12$, where h is the length of preimage interval. For instance, if X is normally distributed the correction to the observed variance is $-8.33 \cdot 10^{-2d}$.

Appendix A

Implementation of the EM Algorithm

A.1 Implementation in R

```
em <- function(mu,sigma,lower, upper, iterations, print){

n<-length(lower)
expectedvalues<-rep(0,n)
expectedvaluessq<-rep(0,n)

for ( j in 1:iterations){

#Probability of X_i being in [L_i,U_i]
if (j==1){
prob<-pnorm(upper,mu,sigma)-pnorm(lower,mu,sigma)
loglikelihoodvalue<-sum(log(prob))
}

#Expected value, E_(mu,sigma^2)(X|X\in [L_i,U_i])
for (i in 1:n){
expectedvalues[i] <-
mu-prob[i]^(-1)*sigma^2/sqrt(2*pi*sigma^2)
*(exp(-(upper[i]-mu)^2/(2*sigma^2))-exp(-(lower[i]-mu)^2/(2*sigma^2)))
}

#Expected value, E_(mu,sigma^2)(X^2|X\in [L_i,U_i])
for (i in 1:n){
expectedvaluessq[i] <-
```

```
mu^2+sigma^2-prob[i]^(-1)*sigma^2/sqrt(2*pi*sigma^2)*
((upper[i]+mu)*exp(-(upper[i]-mu)^2/(2*sigma^2))
-(lower[i]+mu)*exp(-(lower[i]-mu)^2/(2*sigma^2)))
}

#Find new parameters by minimizing the expected loglikelihood
mu<-mean(expectedvalues)
sigma<-sqrt(1/n*sum((expectedvalues-mu)^2+ expectedvaluessq
-expectedvalues^2))

#Check that the loglikelihood function doesn't increase
oldloglikelihoodvalue<-loglikelihoodvalue
#Probability of X being in [L,U]
prob<-pnorm(upper,mu,sigma)-pnorm(lower,mu,sigma)
loglikelihoodvalue<-sum(log(prob))
if (oldloglikelihoodvalue>loglikelihoodvalue){print("ERROR")}
#Print value of loglikelihood function
if (print==TRUE){
print("Iteration:")
print(j)
if (j==1){
print("Loglike starting point")
print(oldloglikelihoodvalue)
}
print("value incomplete loglikelihood")
print(loglikelihoodvalue)
print("Difference loglike")
print(-oldloglikelihoodvalue+loglikelihoodvalue)
print("Mean")
print(mu)
print("Std. deviation")
print(sigma)
}
}
#Return estimate of parameters
return(c(mu,sigma))
}


##Variance estimate
j<-1
cov<-matrix(rep(0,4),2,2)
for (j in 1:600){
parmu<--(dnorm(interval[j,4],mu,sigma)-dnorm(interval[j,3],mu,sigma))
        /(pnorm(interval[j,4],mu,sigma)-pnorm(interval[j,3],mu,sigma))
parsigma<- -1/(2*sigma^2)*((interval[j,4]-mu)*dnorm(interval[j,4],mu,sigma)
-(interval[j,3]-mu)*dnorm(interval[j,3],mu,sigma))
/(pnorm(interval[j,4],mu,sigma)-pnorm(interval[j,3],mu,sigma))
cov<-cov+t(matrix(c(parmu,parsigma),1,2))%*%matrix(c(parmu,parsigma),1,2)
}
solve(cov)
sqrt(solve(cov)[1,1])
sqrt(solve(cov)[2,2])
```

A.2 Implementation in Visual Basic for Excel

```
'This method extracts the values of the three first colums in the sheet
'and uses these for observations, lower bound and upper bound. Based on this
'the MLE is calculated with the EM algorithm. The ML estimate of the mean value is
'given in cell f3 and the ML estimate of the standard deviation is given in cell f4.
'The number of iterations is set to five but can be changed if necessary.
Public Sub CalcEMOnColumnAAndB()


    Dim boundsLower() As Double
    Dim boundsUpper() As Double
    Dim observations() As Double
    observations = GetNumbersInColumn("a1")
    boundsLower = GetNumbersInColumn("b1")
    boundsUpper = GetNumbersInColumn("c1")

    Dim mu As Double
    Dim sigma As Double
    Dim iterations As Integer
    mu = Application.WorksheetFunction.Average(observations)
    sigma = Application.WorksheetFunction.StDev(observations)
    iterations = 5


    'Display an error if the arrays are not of equal length
    If (UBound(boundsLower) <> UBound(boundsUpper)) Then
        MsgBox "The length of the data in column A and B must be of equal length."
        Exit Sub
    End If

    Call em(mu, sigma, boundsLower, boundsUpper, iterations)

    Range("f3").Select
    Selection.Value = mu
    Range("f4").Select
    Selection.Value = sigma


End Sub

Private Sub em(ByRef mu As Double, ByRef sigma As Double,
 boundsLower() As Double, boundsUpper() As Double, ByRef iterations As Integer)

Dim dataLength As Integer
dataLength = UBound(boundsLower)

For iterationCount = 1 To iterations

    Dim prob() As Double
    ReDim prob(1 To dataLength)
    For i = 1 To dataLength
        prob(i) =
Application.WorksheetFunction.NormDist(boundsUpper(i),mu,sigma,True)
    -Application.WorksheetFunction.NormDist(boundsLower(i),mu,sigma,True)
    Next i
```

```
    Dim simpleExpectedValues() As Double
    ReDim simpleExpectedValues(1 To dataLength)
    For i = 1 To dataLength
        simpleExpectedValues(i) = mu-sigma^2*prob(i)^(-1)*
       (Application.WorksheetFunction.NormDist(boundsUpper(i),mu,sigma,False)
       -Application.WorksheetFunction.NormDist(boundsLower(i),mu,sigma,False))
    Next i

    Dim simpleExpectedValuesSq() As Double
    ReDim simpleExpectedValuesSq(1 To dataLength)
    For i = 1 To dataLength
        simpleExpectedValuesSq(i) = mu^2+sigma^2-sigma^2*prob(i)^(-1)*
  ((boundsUpper(i) + mu)*
  Application.WorksheetFunction.NormDist(boundsUpper(i),mu,sigma,False)
  -(boundsLower(i) + mu)
        *Application.WorksheetFunction.NormDist(boundsLower(i),mu,sigma,False))
    Next i

    mu = Application.WorksheetFunction.Average(simpleExpectedValues)

    Dim tmpSum As Double
    tmpSum = 0
    For i = 1 To dataLength
        tmpSum = tmpSum + (simpleExpectedValues(i) - mu) ^ 2
 + simpleExpectedValuesSq(i) - simpleExpectedValues(i) ^ 2
    Next i
    sigma = Sqr((1 / dataLength) * tmpSum)

Next iterationCount

End Sub

Private Function GetNumbersInColumn(topColumnCell As String) As Double()

    Dim result() As Double
    Dim elementsInColumnCount As Integer
    elementsInColumnCount = CountElementsInColumn(topColumnCell)

    ReDim result(1 To elementsInColumnCount)
    Range(topColumnCell).Select

    For i = 1 To elementsInColumnCount
        result(i) = Selection.Offset(i - 1, 0).Value
    Next i

    GetNumbersInColumn = result

End Function


Private Function CountElementsInColumn(topColumnCell As String) As Integer

    Dim elementsInColumnCount As Integer

    Range(topColumnCell).Select
    Do Until Selection.Value = ""
        elementsInColumnCount = elementsInColumnCount + 1
```

```
        Selection.Offset(1, 0).Select
    Loop

    CountElementsInColumn = elementsInColumnCount

End Function
```

Appendix B

Data and Preimages

d	obs	L	U	d	obs	L	U	d	obs	L	U
1	7.092	7.090986	7.092440	2	7.106	7.10505	7.10650	3	7.121	7.11955	7.12100
1	7.104	7.102618	7.104072	2	7.123	7.12245	7.12390	3	7.115	7.11375	7.11520
1	7.111	7.109888	7.111342	2	7.106	7.10505	7.10650	3	7.106	7.10505	7.10650
1	7.104	7.102618	7.104072	2	7.103	7.10215	7.10360	3	7.110	7.10940	7.11085
1	7.102	7.101164	7.102618	2	7.090	7.08910	7.09055	3	7.103	7.10215	7.10360
1	7.106	7.105526	7.106980	2	7.107	7.10650	7.10795	3	7.103	7.10215	7.10360
1	7.099	7.098256	7.099710	2	7.113	7.11230	7.11375	3	7.107	7.10650	7.10795
1	7.099	7.098256	7.099710	2	7.112	7.11085	7.11230	3	7.093	7.09200	7.09345
1	7.111	7.109888	7.111342	2	7.107	7.10650	7.10795	3	7.110	7.10940	7.11085
1	7.088	7.086624	7.088078	2	7.100	7.09925	7.10070	3	7.106	7.10505	7.10650
1	7.115	7.114250	7.115704	2	7.089	7.08765	7.08910	3	7.106	7.10505	7.10650
1	7.108	7.106980	7.108434	2	7.097	7.09635	7.09780	3	7.100	7.09925	7.10070
1	7.108	7.106980	7.108434	2	7.090	7.08910	7.09055	3	7.099	7.09780	7.09925
1	7.106	7.105526	7.106980	2	7.099	7.09780	7.09925	3	7.106	7.10505	7.10650
1	7.089	7.088078	7.089532	2	7.107	7.10650	7.10795	3	7.118	7.11665	7.11810
1	7.105	7.104072	7.105526	2	7.110	7.10940	7.11085	3	7.116	7.11520	7.11665
1	7.108	7.106980	7.108434	2	7.092	7.09055	7.09200	3	7.106	7.10505	7.10650
1	7.111	7.109888	7.111342	2	7.109	7.10795	7.10940	3	7.106	7.10505	7.10650
1	7.099	7.098256	7.099710	2	7.092	7.09055	7.09200	3	7.113	7.11230	7.11375
1	7.093	7.092440	7.093894	2	7.119	7.11810	7.11955	3	7.103	7.10215	7.10360
1	7.109	7.108434	7.109888	2	7.090	7.08910	7.09055	3	7.113	7.11230	7.11375
1	7.118	7.117158	7.118612	2	7.107	7.10650	7.10795	3	7.110	7.10940	7.11085
1	7.104	7.102618	7.104072	2	7.119	7.11810	7.11955	3	7.110	7.10940	7.11085
1	7.102	7.101164	7.102618	2	7.118	7.11665	7.11810	3	7.115	7.11375	7.11520
1	7.108	7.106980	7.108434	2	7.106	7.10505	7.10650	3	7.105	7.10360	7.10505
1	7.109	7.108434	7.109888	2	7.122	7.12100	7.12245	3	7.102	7.10070	7.10215
1	7.101	7.099710	7.101164	2	7.106	7.10505	7.10650	3	7.102	7.10070	7.10215
1	7.101	7.099710	7.101164	2	7.113	7.11230	7.11375	3	7.100	7.09925	7.10070
1	7.096	7.095348	7.096802	2	7.122	7.12100	7.12245	3	7.099	7.09780	7.09925
1	7.109	7.108434	7.109888	2	7.113	7.11230	7.11375	3	7.100	7.09925	7.10070
1	7.102	7.101164	7.102618	2	7.109	7.10795	7.10940	3	7.107	7.10650	7.10795

```
1 7.102 7.101164 7.102618    2 7.103 7.10215 7.10360    3 7.105 7.10360 7.10505
1 7.098 7.096802 7.098256    2 7.109 7.10795 7.10940    3 7.109 7.10795 7.10940
1 7.101 7.099710 7.101164    2 7.105 7.10360 7.10505    3 7.113 7.11230 7.11375
1 7.106 7.105526 7.106980    2 7.099 7.09780 7.09925    3 7.093 7.09200 7.09345
1 7.111 7.109888 7.111342    2 7.106 7.10505 7.10650    3 7.107 7.10650 7.10795
1 7.099 7.098256 7.099710    2 7.103 7.10215 7.10360    3 7.119 7.11810 7.11955
1 7.105 7.104072 7.105526    2 7.106 7.10505 7.10650    3 7.102 7.10070 7.10215
1 7.112 7.111342 7.112796    2 7.107 7.10650 7.10795    3 7.107 7.10650 7.10795
1 7.109 7.108434 7.109888    2 7.100 7.09925 7.10070    3 7.106 7.10505 7.10650
1 7.106 7.105526 7.106980    2 7.107 7.10650 7.10795    3 7.121 7.11955 7.12100
1 7.112 7.111342 7.112796    2 7.097 7.09635 7.09780    3 7.110 7.10940 7.11085
1 7.099 7.098256 7.099710    2 7.099 7.09780 7.09925    3 7.110 7.10940 7.11085
1 7.104 7.102618 7.104072    2 7.121 7.11955 7.12100    3 7.100 7.09925 7.10070
1 7.106 7.105526 7.106980    2 7.112 7.11085 7.11230    3 7.113 7.11230 7.11375
1 7.102 7.101164 7.102618    2 7.123 7.12245 7.12390    3 7.115 7.11375 7.11520
1 7.096 7.095348 7.096802    2 7.096 7.09490 7.09635    3 7.100 7.09925 7.10070
1 7.104 7.102618 7.104072    2 7.107 7.10650 7.10795    3 7.106 7.10505 7.10650
1 7.101 7.099710 7.101164    2 7.109 7.10795 7.10940    3 7.107 7.10650 7.10795
1 7.104 7.102618 7.104072    2 7.110 7.10940 7.11085    3 7.109 7.10795 7.10940
```

d	obs	L	U		d	obs	L	U		d	obs	L	U
4	7.10900	7.10795	7.10940	5	7.1	7.09890	7.10035	6		7.09900	7.09780	7.09925	
4	7.10700	7.10650	7.10795	5	7.10600	7.10470	7.10615	6		7.10000	7.09925	7.10070	
4	7.09900	7.09780	7.09925	5	7.10400	7.10325	7.10470	6		7.10500	7.10360	7.10505	
4	7.10900	7.10795	7.10940	5	7.09600	7.09455	7.09600	6		7.10900	7.10795	7.10940	
4	7.10600	7.10505	7.10650	5	7.09700	7.09600	7.09745	6		7.10700	7.10650	7.10795	
4	7.11000	7.10940	7.11085	5	7.09400	7.09310	7.09455	6		7.10300	7.10215	7.10360	
4	7.10300	7.10215	7.10360	5	7.11400	7.11340	7.11485	6		7.09600	7.09490	7.09635	
4	7.09900	7.09780	7.09925	5	7.10600	7.10470	7.10615	6		7.11300	7.11230	7.11375	
4	7.10700	7.10650	7.10795	5	7.10400	7.10325	7.10470	6		7.10200	7.10070	7.10215	
4	7.10700	7.10650	7.10795	5	7.11900	7.11775	7.11920	6		7.10200	7.10070	7.10215	
4	7.10900	7.10795	7.10940	5	7.08700	7.08585	7.08730	6		7.10900	7.10795	7.10940	
4	7.11300	7.11230	7.11375	5	7.11100	7.11050	7.11195	6		7.10700	7.10650	7.10795	
4	7.09000	7.08910	7.09055	5	7.10700	7.10615	7.10760	6		7.11000	7.10940	7.11085	
4	7.10000	7.09925	7.10070	5	7.10100	7.10035	7.10180	6		7.10300	7.10215	7.10360	
4	7.08700	7.08620	7.08765	5	7.09800	7.09745	7.09890	6		7.10700	7.10650	7.10795	
4	7.10200	7.10070	7.10215	5	7.10000	7.09890	7.10035	6		7.10200	7.10070	7.10215	
4	7.10300	7.10215	7.10360	5	7.10700	7.10615	7.10760	6		7.11600	7.11520	7.11665	
4	7.10300	7.10215	7.10360	5	7.09800	7.09745	7.09890	6		7.11000	7.10940	7.11085	
4	7.10000	7.09925	7.10070	5	7.11700	7.11630	7.11775	6		7.10200	7.10070	7.10215	
4	7.10900	7.10795	7.10940	5	7.11600	7.11485	7.11630	6		7.10600	7.10505	7.10650	
4	7.10600	7.10505	7.10650	5	7.10900	7.10760	7.10905	6		7.10200	7.10070	7.10215	
4	7.11200	7.11085	7.11230	5	7.10900	7.10760	7.10905	6		7.09700	7.09635	7.09780	
4	7.10500	7.10360	7.10505	5	7.10100	7.10035	7.10180	6		7.10500	7.10360	7.10505	
4	7.10200	7.10070	7.10215	5	7.10600	7.10470	7.10615	6		7.09300	7.09200	7.09345	
4	7.09900	7.09780	7.09925	5	7.11100	7.11050	7.11195	6		7.11000	7.10940	7.11085	
4	7.08900	7.08765	7.08910	5	7.10600	7.10470	7.10615	6		7.10200	7.10070	7.10215	
4	7.11200	7.11085	7.11230	5	7.09700	7.09600	7.09745	6		7.10300	7.10215	7.10360	
4	7.10500	7.10360	7.10505	5	7.10100	7.10035	7.10180	6		7.09700	7.09635	7.09780	
4	7.10300	7.10215	7.10360	5	7.09800	7.09745	7.09890	6		7.11500	7.11375	7.11520	
4	7.10700	7.10650	7.10795	5	7.10400	7.10325	7.10470	6		7.10300	7.10215	7.10360	
4	7.11600	7.11520	7.11665	5	7.10300	7.10180	7.10325	6		7.09600	7.09490	7.09635	
4	7.09700	7.09635	7.09780	5	7.10100	7.10035	7.10180	6		7.09700	7.09635	7.09780	
4	7.08100	7.08040	7.08185	5	7.11900	7.11775	7.11920	6		7.10900	7.10795	7.10940	
4	7.10900	7.10795	7.10940	5	7.10600	7.10470	7.10615	6		7.10700	7.10650	7.10795	
4	7.09900	7.09780	7.09925	5	7.11000	7.10905	7.11050	6		7.08300	7.08185	7.08330	
4	7.11200	7.11085	7.11230	5	7.10700	7.10615	7.10760	6		7.09000	7.08910	7.09055	
4	7.10700	7.10650	7.10795	5	7.10400	7.10325	7.10470	6		7.11500	7.11375	7.11520	
4	7.10300	7.10215	7.10360	5	7.09400	7.09310	7.09455	6		7.11300	7.11230	7.11375	
4	7.10300	7.10215	7.10360	5	7.10100	7.10035	7.10180	6		7.09700	7.09635	7.09780	
4	7.10900	7.10795	7.10940	5	7.10900	7.10760	7.10905	6		7.10500	7.10360	7.10505	
4	7.09300	7.09200	7.09345	5	7.10300	7.10180	7.10325	6		7.10000	7.09925	7.10070	
4	7.09900	7.09780	7.09925	5	7.10900	7.10760	7.10905	6		7.11600	7.11520	7.11665	
4	7.10700	7.10650	7.10795	5	7.10900	7.10760	7.10905	6		7.11300	7.11230	7.11375	
4	7.10700	7.10650	7.10795	5	7.09600	7.09455	7.09600	6		7.10600	7.10505	7.10650	
4	7.09900	7.09780	7.09925	5	7.10400	7.10325	7.10470	6		7.10600	7.10505	7.10650	
4	7.09700	7.09635	7.09780	5	7.09800	7.09745	7.09890	6		7.09700	7.09635	7.09780	
4	7.11200	7.11085	7.11230	5	7.12500	7.12355	7.12500	6		7.10700	7.10650	7.10795	
4	7.09700	7.09635	7.09780	5	7.10400	7.10325	7.10470	6		7.11800	7.11665	7.11810	
4	7.10300	7.10215	7.10360	5	7.10600	7.10470	7.10615	6		7.10300	7.10215	7.10360	
4	7.10700	7.10650	7.10795	5	7.11900	7.11775	7.11920	6		7.10600	7.10505	7.10650	

d	obs	L	U	d	obs	L	U	d	obs	L	U
7	7.10000	7.09860	7.10005	8	7.10100	7.10035	7.10180	9	7.10100	7.09995	7.10140
7	7.10200	7.10150	7.10295	8	7.11600	7.11485	7.11630	9	7.10500	7.10430	7.10575
7	7.10100	7.10005	7.10150	8	7.09600	7.09455	7.09600	9	7.10200	7.10140	7.10285
7	7.10200	7.10150	7.10295	8	7.10100	7.10035	7.10180	9	7.11800	7.11735	7.11880
7	7.11700	7.11600	7.11745	8	7.11700	7.11630	7.11775	9	7.10800	7.10720	7.10865
7	7.11400	7.11310	7.11455	8	7.10100	7.10035	7.10180	9	7.09700	7.09560	7.09705
7	7.11600	7.11455	7.11600	8	7.10000	7.09890	7.10035	9	7.11300	7.11155	7.11300
7	7.09800	7.09715	7.09860	8	7.09700	7.09600	7.09745	9	7.10200	7.10140	7.10285
7	7.09800	7.09715	7.09860	8	7.10300	7.10180	7.10325	9	7.10100	7.09995	7.10140
7	7.11300	7.11165	7.11310	8	7.09600	7.09455	7.09600	9	7.11300	7.11155	7.11300
7	7.10800	7.10730	7.10875	8	7.11000	7.10905	7.11050	9	7.11000	7.10865	7.11010
7	7.10200	7.10150	7.10295	8	7.10700	7.10615	7.10760	9	7.09800	7.09705	7.09850
7	7.11300	7.11165	7.11310	8	7.08400	7.08295	7.08440	9	7.09100	7.08980	7.09125
7	7.10200	7.10150	7.10295	8	7.10900	7.10760	7.10905	9	7.10500	7.10430	7.10575
7	7.11000	7.10875	7.11020	8	7.11100	7.11050	7.11195	9	7.10800	7.10720	7.10865
7	7.11300	7.11165	7.11310	8	7.10400	7.10325	7.10470	9	7.11100	7.11010	7.11155
7	7.10400	7.10295	7.10440	8	7.11000	7.10905	7.11050	9	7.11000	7.10865	7.11010
7	7.10800	7.10730	7.10875	8	7.10600	7.10470	7.10615	9	7.12000	7.11880	7.12025
7	7.09500	7.09425	7.09570	8	7.10600	7.10470	7.10615	9	7.10800	7.10720	7.10865
7	7.08800	7.08700	7.08845	8	7.09000	7.08875	7.09020	9	7.09900	7.09850	7.09995
7	7.10800	7.10730	7.10875	8	7.10400	7.10325	7.10470	9	7.12000	7.11880	7.12025
7	7.08400	7.08265	7.08410	8	7.10100	7.10035	7.10180	9	7.10700	7.10575	7.10720
7	7.10100	7.10005	7.10150	8	7.10600	7.10470	7.10615	9	7.09800	7.09705	7.09850
7	7.09200	7.09135	7.09280	8	7.10700	7.10615	7.10760	9	7.10200	7.10140	7.10285
7	7.11100	7.11020	7.11165	8	7.10400	7.10325	7.10470	9	7.10400	7.10285	7.10430
7	7.09700	7.09570	7.09715	8	7.11100	7.11050	7.11195	9	7.10500	7.10430	7.10575
7	7.10200	7.10150	7.10295	8	7.11600	7.11485	7.11630	9	7.10700	7.10575	7.10720
7	7.10500	7.10440	7.10585	8	7.10700	7.10615	7.10760	9	7.12100	7.12025	7.12170
7	7.10100	7.10005	7.10150	8	7.11100	7.11050	7.11195	9	7.10200	7.10140	7.10285
7	7.10400	7.10295	7.10440	8	7.11100	7.11050	7.11195	9	7.10800	7.10720	7.10865
7	7.11300	7.11165	7.11310	8	7.10100	7.10035	7.10180	9	7.11300	7.11155	7.11300
7	7.11300	7.11165	7.11310	8	7.10300	7.10180	7.10325	9	7.10200	7.10140	7.10285
7	7.11300	7.11165	7.11310	8	7.10400	7.10325	7.10470	9	7.09800	7.09705	7.09850
7	7.11400	7.11310	7.11455	8	7.10600	7.10470	7.10615	9	7.09500	7.09415	7.09560
7	7.11300	7.11165	7.11310	8	7.10100	7.10035	7.10180	9	7.11500	7.11445	7.11590
7	7.10400	7.10295	7.10440	8	7.10400	7.10325	7.10470	9	7.09900	7.09850	7.09995
7	7.10500	7.10440	7.10585	8	7.10400	7.10325	7.10470	9	7.10400	7.10285	7.10430
7	7.10100	7.10005	7.10150	8	7.09800	7.09745	7.09890	9	7.10200	7.10140	7.10285
7	7.10800	7.10730	7.10875	8	7.09700	7.09600	7.09745	9	7.09700	7.09560	7.09705
7	7.10200	7.10150	7.10295	8	7.10700	7.10615	7.10760	9	7.09800	7.09705	7.09850
7	7.11300	7.11165	7.11310	8	7.10300	7.10180	7.10325	9	7.10400	7.10285	7.10430
7	7.10200	7.10150	7.10295	8	7.11000	7.10905	7.11050	9	7.09900	7.09850	7.09995
7	7.11300	7.11165	7.11310	8	7.10300	7.10180	7.10325	9	7.11100	7.11010	7.11155
7	7.10700	7.10585	7.10730	8	7.09700	7.09600	7.09745	9	7.10800	7.10720	7.10865
7	7.10700	7.10585	7.10730	8	7.10300	7.10180	7.10325	9	7.10800	7.10720	7.10865
7	7.10200	7.10150	7.10295	8	7.11900	7.11775	7.11920	9	7.10100	7.09995	7.10140
7	7.09400	7.09280	7.09425	8	7.10900	7.10760	7.10905	9	7.09900	7.09850	7.09995
7	7.10400	7.10295	7.10440	8	7.11700	7.11630	7.11775	9	7.10500	7.10430	7.10575
7	7.11100	7.11020	7.11165	8	7.10100	7.10035	7.10180	9	7.09700	7.09560	7.09705
7	7.10000	7.09860	7.10005	8	7.10300	7.10180	7.10325	9	7.10800	7.10720	7.10865

d	obs	L	U	d	obs	L	U	d	obs	L	U
10	7.10900	7.10795	7.10940	11	7.11300	7.11230	7.11375	12	7.08900	7.08765	7.08910
10	7.09900	7.09780	7.09925	11	7.10300	7.10215	7.10360	12	7.09200	7.09055	7.09200
10	7.10000	7.09925	7.10070	11	7.11200	7.11085	7.11230	12	7.10900	7.10795	7.10940
10	7.09900	7.09780	7.09925	11	7.10900	7.10795	7.10940	12	7.11000	7.10940	7.11085
10	7.10000	7.09925	7.10070	11	7.11000	7.10940	7.11085	12	7.11000	7.10940	7.11085
10	7.10900	7.10795	7.10940	11	7.10000	7.09925	7.10070	12	7.08700	7.08620	7.08765
10	7.10900	7.10795	7.10940	11	7.09900	7.09780	7.09925	12	7.11000	7.10940	7.11085
10	7.10900	7.10795	7.10940	11	7.10600	7.10505	7.10650	12	7.10500	7.10360	7.10505
10	7.09000	7.08910	7.09055	11	7.11000	7.10940	7.11085	12	7.09200	7.09055	7.09200
10	7.09600	7.09490	7.09635	11	7.12100	7.11955	7.12100	12	7.10700	7.10650	7.10795
10	7.11000	7.10940	7.11085	11	7.10900	7.10795	7.10940	12	7.11000	7.10940	7.11085
10	7.10900	7.10795	7.10940	11	7.10900	7.10795	7.10940	12	7.10000	7.09925	7.10070
10	7.10600	7.10505	7.10650	11	7.11900	7.11810	7.11955	12	7.09200	7.09055	7.09200
10	7.10900	7.10795	7.10940	11	7.11300	7.11230	7.11375	12	7.10000	7.09925	7.10070
10	7.10000	7.09925	7.10070	11	7.10500	7.10360	7.10505	12	7.09400	7.09345	7.09490
10	7.10700	7.10650	7.10795	11	7.11800	7.11665	7.11810	12	7.10500	7.10360	7.10505
10	7.10700	7.10650	7.10795	11	7.11200	7.11085	7.11230	12	7.10600	7.10505	7.10650
10	7.11000	7.10940	7.11085	11	7.10000	7.09925	7.10070	12	7.08900	7.08765	7.08910
10	7.11500	7.11375	7.11520	11	7.11500	7.11375	7.11520	12	7.08600	7.08475	7.08620
10	7.08400	7.08330	7.08475	11	7.10300	7.10215	7.10360	12	7.10200	7.10070	7.10215
10	7.10700	7.10650	7.10795	11	7.10700	7.10650	7.10795	12	7.10300	7.10215	7.10360
10	7.10000	7.09925	7.10070	11	7.11200	7.11085	7.11230	12	7.10000	7.09925	7.10070
10	7.09000	7.08910	7.09055	11	7.09300	7.09200	7.09345	12	7.10200	7.10070	7.10215
10	7.11300	7.11230	7.11375	11	7.10500	7.10360	7.10505	12	7.09200	7.09055	7.09200
10	7.11800	7.11665	7.11810	11	7.10700	7.10650	7.10795	12	7.09300	7.09200	7.09345
10	7.10000	7.09925	7.10070	11	7.11000	7.10940	7.11085	12	7.08900	7.08765	7.08910
10	7.09200	7.09055	7.09200	11	7.10600	7.10505	7.10650	12	7.09600	7.09490	7.09635
10	7.10700	7.10650	7.10795	11	7.09700	7.09635	7.09780	12	7.10700	7.10650	7.10795
10	7.11200	7.11085	7.11230	11	7.10700	7.10650	7.10795	12	7.10300	7.10215	7.10360
10	7.10900	7.10795	7.10940	11	7.10300	7.10215	7.10360	12	7.08600	7.08475	7.08620
10	7.10700	7.10650	7.10795	11	7.10900	7.10795	7.10940	12	7.11200	7.11085	7.11230
10	7.10300	7.10215	7.10360	11	7.10600	7.10505	7.10650	12	7.10900	7.10795	7.10940
10	7.10500	7.10360	7.10505	11	7.11000	7.10940	7.11085	12	7.11000	7.10940	7.11085
10	7.10500	7.10360	7.10505	11	7.10600	7.10505	7.10650	12	7.09700	7.09635	7.09780
10	7.10200	7.10070	7.10215	11	7.09200	7.09055	7.09200	12	7.11000	7.10940	7.11085
10	7.10600	7.10505	7.10650	11	7.10700	7.10650	7.10795	12	7.09400	7.09345	7.09490
10	7.10500	7.10360	7.10505	11	7.11300	7.11230	7.11375	12	7.09000	7.08910	7.09055
10	7.09600	7.09490	7.09635	11	7.10300	7.10215	7.10360	12	7.08900	7.08765	7.08910
10	7.10600	7.10505	7.10650	11	7.11800	7.11665	7.11810	12	7.10600	7.10505	7.10650
10	7.10900	7.10795	7.10940	11	7.10600	7.10505	7.10650	12	7.11500	7.11375	7.11520
10	7.09700	7.09635	7.09780	11	7.10900	7.10795	7.10940	12	7.11500	7.11375	7.11520
10	7.09900	7.09780	7.09925	11	7.09700	7.09635	7.09780	12	7.10700	7.10650	7.10795
10	7.10300	7.10215	7.10360	11	7.10700	7.10650	7.10795	12	7.10600	7.10505	7.10650
10	7.10200	7.10070	7.10215	11	7.11000	7.10940	7.11085	12	7.10900	7.10795	7.10940
10	7.10900	7.10795	7.10940	11	7.09700	7.09635	7.09780	12	7.09300	7.09200	7.09345
10	7.11000	7.10940	7.11085	11	7.11300	7.11230	7.11375	12	7.11800	7.11665	7.11810
10	7.10500	7.10360	7.10505	11	7.10700	7.10650	7.10795	12	7.10700	7.10650	7.10795
10	7.10700	7.10650	7.10795	11	7.11300	7.11230	7.11375	12	7.11500	7.11375	7.11520
10	7.09300	7.09200	7.09345	11	7.11000	7.10940	7.11085	12	7.10300	7.10215	7.10360
10	7.10500	7.10360	7.10505	11	7.11000	7.10940	7.11085	12	7.11800	7.11665	7.11810

Bibliography

Anderson, M. J. and Whitcomb, P. J. (2007), "Using Graphical Diagnostic to Deal with Bad Data," *Quality Engineering*, 19, 111–118.

Anderson, T. W. (1958), *An introduction to multivariate statistical analysis*, New York: John Wiley & Sons.

Bischak, D. P. and Trietsch, D. (2007), "The Rate of False Signals in X-bar Control Charts with Estimated Limits," *Journal of Quality Techology*, 39, 54–65.

Bisgaard, S. and Kulahci, M. M. (2005), "Checking Process Stability with the Variogram," *Quality Engineering*, 17, 323–327.

Box, G. and Paniagua-Quinones, C. (2007), "Two Charts: Not One," *Qualilty Engineering*, 19, 93–100.

Box, G. E. P. (1949), "A General Distribution Theory for a Class of Likelihood Criteria," *Biometrika*, 36, 318–335.

— (1979), *Robustness in the Strategy of Scientific Model Building*, New York: Academic Press.

Box, G. E. P., Hunter, J. S., and Hunter, W. G. (2005), *Statistics For Experimenters, 2nd ed.*, Hoboken, NJ: John Wiley & Sons.

Braun, W. J. and Park, D. (2008), "Estimation of σ for Individuals Charts," *Journal of Quality Technology*, 40, 332–344.

Carter, E. M. and Srivastava, M. S. (1977), "Monotonicity of the Power Functions of Modified Likelihood Ratio Criterion for the Homogeneity of Variances and of the Sphericity Test," *Journal of Multivariate Analysis*, 7, 229–233.

Castagliola, P. (2001), "How to monitor capability index C_p using EWMA," *International Journal of Reliability, Quality, and Safety Engineering*, 8, 191–204.

Castagliola, P. and Vännman, K. (2007), "Monitoring capability indices using an EWMA approach," *Quality and Reliability Engineering International*, 23, 769–790.

— (2008), "Average Run Length When Monitoring Capability Indices Using EWMA," *Quality and Reliability Engineering International*, 24, 941–955.

Chakraborti, S., Human, S. W., and Graham, M. A. (2009), "Phase I Statistical Process Control: An overview and Some Results," *Quality Engineering*, 21, 52–62.

Chan, L. K., Cheng, W., and Spiring, F. A. (1988), "A New Meassure of Process Capability: C_{pm}," *Journal of Quality Technology*, 20, 162–175.

Chou, C.-Y., Chen, C.-H., and Liu, H.-R. (2005), "Acceptance Control Charts for Non-normal Data," *Journal of Applied Statistics*, 32, 25–36.

Chou, Y.-M., O, D. B., Salvador, A., and Borrega, A. (1990), "Lower Confidence Limits on Process Capability Indices," *Journal of Quality Technology*, 22, 223–229.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39, 1–38.

Freund, R. A. (1957), "Acceptance control charts," *Industrial Quality Control*, 14, 13–23.

Gleser, L. J. (1966), "A Note on the Sphericity Test," *The Annals of Mathematical Statistics*, 37, 464–467.

Grant, E. (1946), *Statistical Quality Control*, New York: McGraw-Hill Book Company.

Grant, E. and Leavenworth, R. (1989), *Statistical Quality Control, 6th ed.*, New York: McGraw-Hill Book Company.

Hill, D. (1956), "Modified Control Limits," *Applied Statistics*, 5, 12–19.

Holmes, D. and Mergen, A. (2000), "Exponentiallly Weighted Moving Average Charts," *Quality Reliability Engineering International*, 16, 139–142.

Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall, W. H. (2006), "Effect of Parameter Estimation on Control Chart Properties: A Literature Review," *Journal of Quality Techology*, 38, 349–369.

Jones, P. N. and Mclachlan, G. J. (1992), "Improving the Convergence rate of the EM algorithm for a mixture model fitted to grouped truncated data," *Journal of Statistical Computation and Simulation*, 43, 31–44.

Jones-Farmer, L. A., Jordan, V., and Champ, C. W. (2009), "Distribution-Free Phase Control Charts for Subgroup Location," *Journal of Quality Technology*, 41, 304–316.

Kotz, S. and Johnson, N. (2002), "Process Capability – A Review, 1992–2000," *Journal of Quality Technology*, 34, 2–18.

Kotz, S. and Lovelace, C. R. (1998), *Process Capability Indicies in Theory and Practice*, New York: Oxford University Press Inc.

Mast, J. D. and Trip, A. (2009), "Exploratory Data Analysis in Quality-Improvement Projects," *Journal of Quality Technology*, 39, 301–311.

Mauchly, J. W. (1940), "Significance Test for Sphericity of a Normal n-Variate Distribution," *The Annals of Mathematical Statistics*, 11, 204–209.

McLachlan, G. J. and Krishnan, T. (1997), *The EM Algorithm and Extensions, Wiley Series in Probability and Statistis*, New York: John Wiley & Sons, Inc.

Montgomery, D. C. (2005), *Introduction to statistical quality control, 5th ed.*, Hoboken, NJ: John Wiley & Sons.

Muirhead, R. J. (1982), *Aspect of Multivariate Statistical Theory*, New York: John Wiley & Sons.

Nagarsenker, B. N. and Pillai, K. C. S. (1973), "The Distribution of the Sphericity Test Criterion," *Journal of Multivariate Analysis*, 3, 226–235.

Palm, A. (2000), "Controversies and Contradictions in Statistical Process Control: Discussion," *Journal of Quality Technology*, 32, 356–360.

Plante, J.-F. and Windfeldt, G. B. (2009), "Using Predictive Risk for Process Monitoring," *submitted*.

Roes, K. C. B. and Does, R. J. M. M. (1995), "Shewhart-Type Charts in Nonstandard Situations," *Technometrics*, 37, 15–24.

Spiring, F., Bartholomew, L., Smiley, C., and Anthony, Y. (2003), "A Bibliography of Process Capability Papers," *Quality Reliability Engineering International*, 19, 445–460.

Spiring, F. A. (1991), "Assesing process capability in the presences of systematic assignable causes," *Journal of Quality Technology*, 23, 125–134.

Steiner, S. and Wesolowsky, G. (1994), "Simultanious Acceptance Control Charts for Multiple Correlated Characteristics," *International Journal of Production Research*, 32, 531–543.

Steiner, S. H. (1998), "Grouped Data Exponentially Weighted Moving Average Control Charts," *Applied Statistics*, 47, 203–216.

Steiner, S. H., Geyer, P. L., , and Wesolowsky, G. O. (1996), "Shewhart Control Charts to Detect Mean Shifts Based on Grouped Data," *Quality Reliability Engineering International*, 12, 345–353.

Stuart, A. and Ord, J. K. (1998), *Kendall's Advanced Theory of Statistics, Vol. 1, Distribution Theory, 6th ed.*, New York: Oxford University Press.

Sugiura, N. and Nagao, H. (1968), "Unbiasedness of some test criteria for the equality of one or two covariance matrices," *The Annals of Mathematical Statistics*, 39, 1686–1692.

Sullivan, J. H., Woodall, W. H., and Gardner, M. M. (1995), "Discussion of Shewhart-Type Charts in Nonstandard Situations," *Technometrics*, 37, 31–35.

Thyregod, P. and Iwersen, J. (2000), "Controversies and Contradictions in Statistical Process Control: Discussion," *Journal of Quality Technology*, 32, 356–376.

Vännman, K. (1995), "A Unified Approach of Capability Indicies," *Statistica Sinica*, 5, 805–820.

Vining, G. (2009), "Technical Advice: Phase I and Phase II Control Charts," *Quality Engineering*, 21, 478–479.

Wesolowsky, G. O. (1990), "Simultanious Acceptance Control Charts for Two Correlated Processes," *Technometrics*, 32, 43–48.

— (1992), "Simultanious Acceptance Control Charts for Independent Processes," *Applied Statistics*, 41, 147–158.

Windfeldt, G. B. and Bisgaard, S. (2009), "Testing for Sphericity in Phase I Control Chart Applications," *Quality Reliability Engineering International*, 25, 839–849.

Windfeldt, G. B. and Hartvig, N. V. (2010), "Assessing the Impact of Missing Values on Quality Measures of an Industrial Process," *submitted*.

Woodall, W. H. (2000), "Controversies and contradictions in statistical process control," *Journal of Quality Technology*, 32, 341–350.

Woodall, W. H. and Thomas, E. V. (1995), "Statistical Process Control with Several Components of Common Cause Variability," *IIE Transactions*, 27, 757–764.

Wu, Z. (1998), "An Adaptive Acceptance Chart for Tool Wear," *International Journal of Production Research*, 36, 1571–1586.

Summary in English

Paper 1 is aimed at practicians to help them test the assumption that the observations in a sample are independent and identically distributed. An assumption that is essential when using classical Shewhart charts. The test can easily be performed in the control chart setup using the samples gathered here and standard statistical software.

In Paper 2 a new method for process monitoring is introduced. The method uses a statistical model of the quality characteristic and a sliding window of observations to estimate the probability that the next item will not respect the specifications. If the estimated probability exceeds a pre-determined threshold the process will be stopped. The method is flexible, allowing a complexity in modeling that remains invisible to the end user. Furthermore, the method allows to build diagnostic plots based on the parameters estimates that can provide valuable insight into the process. The method is explored numerically and a case study is provided. In Paper 3 the method is explored in a bivariate setting.

Paper 4 is a case study on a problem regarding missing values in an industrial process. The impact of the missing values on the quality measures of the process is assessed. Furthermore, guidelines along with software is provided to handle similar problems.

Summary in Danish

Artikel 1 er målrettet praktikere for at hjælpe dem med at teste antagelsen om observationerne i en stikprøve er uafhængige og identisk fordelte. En antagelse der er fundamental når man benytter klassiske Shewhart kort. Testen kan nemt udføres i kontrol korts opsætningsfasen ved hjælp af de stikprøver der er samlet og standard statistisk software.

I artikel 2 introduceres en ny metode til proces monitorering. Metoden benytter en statistisk model og et glidende vindue af observationer til at estimere sandsynligheden for at det næste emne ikke overholder specifikationerne. Sfremt den estimerede sandsynlighed overstiger en forudbestemt grænseværdi stoppes processen. Metoden er fleksibel og tillader at benytte komplekse modeller som er usynlige for slutbrugeren. Desuden giver metoden mulighed for at lave plots af parameter estimaterne, hvilket kan give værdifuld indsigt i processen. Metoden undersøges numerisk og en case study præsenteres. I artikel 3 undersøges metoden for en bivariat proces.

Artikel 4 er en case study vedrørende manglende værdier i en industriel proces. Betydningen af de manglende værdier for processens kvalitets parametre vurderes. Desuden gives der retningslinier og software til at håndtere lignende problemer.